



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/29, C07K 14/415, C12N 5/10, A01H 5/00, A01N 65/00</b>		<b>A2</b>	(11) International Publication Number: <b>WO 99/38977</b>
(21) International Application Number: <b>PCT/US99/01895</b>		(43) International Publication Date: <b>5 August 1999 (05.08.99)</b>	
(22) International Filing Date: <b>28 January 1999 (28.01.99)</b>		Filed on <b>3 February 1998 (03.02.98)</b> US <b>09/198,119 (CIP)</b> Filed on <b>23 November 1998 (23.11.98)</b>	
(30) Priority Data: 09/018,233 3 February 1998 (03.02.98) US 09/017,816 3 February 1998 (03.02.98) US 09/018,235 3 February 1998 (03.02.98) US 09/017,575 3 February 1998 (03.02.98) US 09/018,227 3 February 1998 (03.02.98) US 09/018,234 3 February 1998 (03.02.98) US 09/198,119 23 November 1998 (23.11.98) US		(71) Applicants (for all designated States except US): MENDEL BIOTECHNOLOGY, INC. [US/US]; 21375 Cabot Boulevard, Hayward, CA 94545 (US). MICHIGAN STATE UNIVERSITY [US/US]; 238 Administration Building, East Lansing, MI 48824-1046 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): STOCKINGER, Eric, J. [US/US]; 1360 Burcham Drive, East Lansing, MI 48823 (US). JAGLO-OTTOSEN, Kirsten [US/US]; 307 South Clemens Avenue, Lansing, MI 48912 (US). ZARKA, Daniel [US/US]; 2101 Barritt Street, Lansing, MI 48912 (US). GILMOUR, Sarah, Jane [GB/US]; 1830 Barnes Road, Leslie, MI 49251 (US). JIANG, Cai-Zhong [CN/US]; 34495 Heathrow Terrace, Fremont, CA 94555 (US). FROMM, Michael [US/US]; 968 Keeler Avenue, Berkeley, CA 94708 (US). THOMASHOW, Michael, F. [US/US]; 805 Southlawn Avenue, East Lansing, MI 48823 (US). (74) Agent: GUERRERO, Karen., J.; MENDEL BIOTECHNOLOGY, INC., 21375 Cabot Boulevard, Hayward, CA 94545 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published Without international search report and to be republished upon receipt of that report.	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications			
US 09/018,233 (CIP) Filed on 3 February 1998 (03.02.98) US 09/017,816 (CIP) Filed on 3 February 1998 (03.02.98) US 09/018,235 (CIP) Filed on 3 February 1998 (03.02.98) US 09/017,575 (CIP) Filed on 3 February 1998 (03.02.98) US 09/018,227 (CIP) Filed on 3 February 1998 (03.02.98) US 09/018,234 (CIP)			
(54) Title: <b>PLANT HAVING ALTERED ENVIRONMENTAL STRESS TOLERANCE</b>			
(57) Abstract			
<p>A transformed plant is provided which comprises one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; a sequence encoding a binding protein capable of binding to the DNA regulatory sequence and inducing expression of the one or more environmental stress tolerance genes; and a recombinant promoter which regulates expression of the gene encoding the binding protein. A method for altering an environmental stress tolerance of a plant is also provided which comprises the steps of transforming a plant with a promoter which regulates expression of at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress tolerance gene through binding of the binding protein to the DNA regulatory sequence.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## PLANT HAVING ALTERED ENVIRONMENTAL STRESS TOLERANCE

### FIELD OF THE INVENTION

5 The present invention relates to the regulatory response of plants to environmental stresses such as cold and to drought. More specifically, the present invention relates to genes which regulate the response of a plant to environmental stresses such as cold or drought and their use to enhance the stress tolerance of recombinant plants into which these genes are introduced.

### 10 BACKGROUND OF THE INVENTION

Environmental factors serve as cues to trigger a number of specific changes in plant growth and development. One such factor is low temperature. Prominent examples of cold-regulated processes include cold acclimation, the increase in freezing tolerance that occurs in response to low non-freezing  
15 temperatures (Guy, C. L., *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**:187-223 (1990)); vernalization, the shortening of time to flowering induced by low temperature (Lang, A., in *Encyclopedia of Plant Physiology*, Vol. 15-1, ed. Ruhland, W. (Springer, Berlin), pp. 1489-1536 (1965)); and stratification, the  
20 breaking of seed dormancy by low temperature (Berry, J. A. and J. K. Raison, in *Encyclopedia of Plant Physiology*, Vol. 12A, eds. Lange, O. L., Nobel, P. S., Osmond, C. B. and Ziegler, H. (Springer, Berlin), pp. 277-338 (1981)). Due to the fundamental nature and agronomic importance of these processes, there is interest in understanding how plants sense and respond to low temperature.  
25 One approach being taken is to determine the signal transduction pathways and regulatory mechanisms involved in cold-regulated gene expression.

Strong evidence exists for calcium having a role in low temperature signal transduction and regulation of at least some COR (cold-regulated) genes. Dhindsa and colleagues (Monroy, A. F., et al, *Plant Physiol.* **102**:1227-  
30 1235 (1993); Monroy, A. F., and R. S., *The Plant Cell*, **7**:321-331 (1995)) have shown that, in alfalfa, calcium chelators and calcium channel blockers prevent low temperature induction of COR genes and that calcium ionophores and calcium channel agonists induce expression of COR genes at normal growth

temperatures. Similarly, Knight et al (The Plant Cell 8:489-503 (1996)) have shown that cold-induced expression of the *Arabidopsis thaliana* COR gene *KIN1* is inhibited by calcium chelators and calcium channel blockers. These results suggest that low temperature triggers an influx of extracellular calcium that activates a signal transduction pathway that induces the expression of COR genes. Consistent with this notion is the finding that low temperature evokes transient increases in cytosolic calcium levels in plants (Knight, M. R. et al, Nature 352:524-526 (1991); Knight, H., et al., The Plant Cell 8:489-503 (1996)). In addition, low temperatures have been shown to stimulate the activity of mechanosensitive calcium-selective cation channels in plants (Ding, J. P. and B. G. Pickard, Plant J. 3:713-720 (1993)).

Recent efforts have led to the identification of a *cis*-acting cold-regulatory element in plants, the C-repeat/DRE (Yamaguchi-Shinozaki, et al., The Plant Cell 6:251-264 (1994); Baker, S. S., et al., Plant. Mol. Biol. 24:701-713 (1994); Jiang, C., et al., Plant Mol. Biol. 30:679-684 (1996)). The element, which has a 5 base pair core sequence for CCGAC, is present once to multiple times in all plant cold-regulated promoters that have been described to date; these include the promoters of the *COR15a* (Baker, S. S., et al, Plant. Mol. Biol. 24:701-713 (1994)), *COR78/RD29A* (Horvath, D. P., et al, Plant Physiol. 103:1047-1053 (1993); Yamaguchi-Shinozaki, K., et al., The Plant Cell 6:251-264 (1994)), *COR6.6* (Wang, H., et al., Plant Mol. Biol. 28:605-617 (1995)) and *KIN1* (Wang, H., et al, Plant Mol. Biol. 28:605-617 (1995)) genes of *Arabidopsis* and the *BN115* gene of *Brassica napus* (White, T. C., et al, Plant Physiol. 106:917-928 (1994)). Deletion analysis of the *Arabidopsis COR15a* gene suggested that the CCGAC sequence, designated the C-repeat, might be part of a *cis*-acting cold-regulatory element (Baker, S. S., et al., Plant Mol. Biol. 24:701-713 (1994)). That this was the case was first demonstrated by Yamaguchi-Shinozaki and Shinozaki (Yamaguchi-Shinozaki, K., et al., The Plant Cell 6:251-264 (1994)) who showed that two of the C-repeat sequences present in the promoter of *COR78/RD29A* induced cold-regulated gene expression when fused to a reporter gene. It was also found that these two elements stimulate transcription in response to dehydration and high salinity and thus, was designated the DRE (dehydration, low temperature and high salt regulatory element). Recent studies by Jiang et al (Jiang, C., et al., Plant Mol.



Biol. 30:679-684 (1996)) indicate that the C-repeats (referred to as low temperature response elements) present in the promoter of the *B. napus* BN115 gene also impart cold-regulated gene expression.

5 U.S. Patents Nos. 5,296,462 and 5,356,816 to Thomashow describe the genes encoding the proteins involved in cold adaptation in *Arabidopsis thaliana*. In particular the DNA encoding the COR15 proteins is described. These proteins are significant in promoting cold tolerance in plants.

10 A need exists for the identification of genes which regulate the expression of cold tolerance genes and drought tolerance genes. A further need exists for DNA constructs useful for introducing these regulatory genes into a plant in order to cause the plant to begin expressing or enhance their expression of native or non-native cold tolerance genes and drought tolerance genes. These and other needs are provided by the present invention.

15

## SUMMARY OF THE INVENTION

DNA in isolated form is provided which includes a sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. In one embodiment, the binding protein is a non-naturally occurring protein formed by combining an amino acid sequence capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence with an amino acid sequence which forms a transcription activation region which regulates expression of one or more environmental stress tolerance genes in a plant by regulating expression of one or more environmental stress tolerance genes when the binding protein binds to the regulatory region.

DNA in isolated form is also provided which includes a promoter and the sequence encoding the binding protein. In one variation, the promoter causes expression of the binding protein in a manner which is different than how the binding protein is expressed in its native state. For example, the promoter may increase the level at which the binding protein is expressed, express the binding protein without being induced by an environmental stress and/or express the binding protein in response to a different form or degree of environmental stress than would otherwise be needed to induce expression of the binding protein. The promoter may also be inducible by an exogenous agent. The promoter can also be selected with regard to the type or types of plant tissues that the binding protein will be expressed as well as when in the plant's life the promoter will function to regulate expression of the binding protein.

A nucleic acid construct capable of transforming a plant is also provided which includes a sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental

stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The nucleic acid construct may be an RNA or DNA construct. Examples of types of constructs include, but are not limited to DNA and RNA viral vectors and plasmids.

A nucleic acid construct capable of transforming a plant is also provided which includes a sequence which when transformed into a plant expresses a binding protein capable of selectively binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant. The binding protein preferably regulates expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

In one variation of the above constructs, the construct also includes a promoter which regulates expression of the binding protein encoding sequence. The promoter may optionally be homologous or heterologous relative to the binding protein encoding sequence. The promoter and binding protein encoding sequence may also optionally be native to the same or a different plant species. In one variation, the promoter causes expression of the binding protein in a manner which is different than how the binding protein is expressed in its native state. For example, the promoter may increase the level at which the binding protein is expressed, express the binding protein without being induced by an environmental stress and/or express the binding protein in response to a different form or degree of environmental stress than would otherwise be needed to induce expression of the binding protein. The promoter may also be inducible by an exogenous agent. The promoter can also be selected with regard to the type or types of plant tissues that the binding protein will be expressed as well as when in the plant's life the promoter will function to regulate expression of the binding protein. For example, flower-, fruit- and seed-specific promoters can be used to regulate the expression of the binding protein in these tissues of the plant, especially when sudden frosts strike in early spring and late fall.

A binding protein in isolated form is also provided which is capable of selectively binding to a DNA regulatory sequence which regulates expression

of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes.

A recombinant binding protein expressed within a plant is also provided which is capable of selectively binding to a DNA regulatory sequence in the plant which regulates expression of one or more environmental stress tolerance genes in the plant. The recombinant binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The recombinant binding protein may be native or non-native to the plant. Further, the recombinant binding protein may be homologous or heterologous relative to the DNA binding protein present in the plant in which the binding protein is expressed.

A transformed cell of an organism is also provided which includes a recombinant sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The transformed cell may be a unicellular organism such as a bacterium, yeast or virus, or from a multicellular organism such as a fungus or a plant.

A transformed cell is also provided which includes a promoter and a sequence encoding a binding protein where at least one of the promoter and sequence under regulatory control of the promoter is recombinant. Optionally, one or both of the promoter and sequence under regulatory control of the promoter is not native to the cell. In one variation, the promoter causes expression of the binding protein in a manner which is different than how the binding protein is expressed in its native state. For example, the promoter may increase the level at which the binding protein is expressed, express the

binding protein without being induced by an environmental stress and/or express the binding protein in response to a different form or degree of environmental stress than would otherwise be needed to induce expression of the binding protein. The promoter may also be inducible by an exogenous agent. The promoter can also be selected with regard to the type or types of plant tissues that the binding protein will be expressed as well as when in the plant's life the promoter will function to regulate expression of the binding protein.

A transformed cell is also provided which includes a recombinant binding protein expressed within the cell which is capable of selectively binding to a DNA regulatory sequence in the plant which regulates expression of one or more environmental stress tolerance genes in the plant. The binding protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in the plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The binding protein may be native or non-native to the cell.

A transformed plant with modified environmental stress tolerance gene expression is also provided. In one embodiment, the transformed plant includes one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and a recombinant sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

In another embodiment, the transformed plant includes one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence; and a recombinant promoter which regulates expression of the sequence encoding the binding protein.

In yet another embodiment, the transformed plant includes one or more environmental stress tolerance genes; a recombinant DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; and a sequence encoding a binding protein capable of selectively binding to the DNA regulatory sequence.

In yet another embodiment, the transformed plant includes at least one recombinant environmental stress tolerance gene; a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and a sequence encoding a binding protein capable of selectively  
5 binding to the DNA regulatory sequence.

In yet another embodiment, the transformed plant includes at least one recombinant environmental stress tolerance gene; a DNA regulatory sequence which regulates expression of the at least one environmental stress tolerance gene; and a recombinant binding protein expressed by the plant which is  
10 capable of selectively binding to the DNA regulatory sequence.

A method for altering an environmental stress tolerance of a plant is also provided. In one embodiment, the method includes transforming a plant with at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental  
15 stress tolerance genes in the plant; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress tolerance gene through binding of the binding protein to the DNA regulatory sequence.

In another embodiment, the method includes transforming a plant with a  
20 promoter which regulates expression of at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant or in specific type or types of plant tissues; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress  
25 tolerance gene through binding of the binding protein to the DNA regulatory sequence.

In another embodiment, the method includes transforming a plant with one or more environmental stress tolerance genes whose expression is regulated by a DNA regulatory sequence; and expressing a binding protein  
30 capable of binding to the DNA regulatory sequence and activating expression of the one or more environmental stress tolerance genes.

According to any one of the above embodiments of the present invention, the binding protein may optionally be selected such that it selectively binds to a member of a class of DNA regulatory sequences which includes the

subsequence CCG or more particularly one of the following subsequences:  
CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG,  
CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG,  
AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG,  
5 CACCG CCCG, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT,  
ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT,  
CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG. The binding  
protein may also be selected such that the binding protein includes an AP2  
domain.

10 In each of the above embodiments, the level of expression of the  
binding protein may be the same or different than the level of expression of the  
binding protein in its native state. Expression of the binding protein in the  
transformed cell may be regulated by a recombinant promoter which may have  
the effect of increasing the level at which the binding protein is expressed,  
15 expressing the binding protein without being induced by an environmental  
stress and/or expressing the binding protein in response to a different form or  
degree of environmental stress than is otherwise needed to induce expression  
of the binding protein. Expression may also be induced by an exogenous  
agent. Expression may also be limited to selected types of plant tissues or  
20 selected periods in the plant's life based on which promoter is used. By  
selecting in what tissues and when in a plant's life the binding protein is  
expressed, in combination with the selecting how the binding protein is  
expressed (level of expression and/or type of environmental or chemical  
induction), an incredible range of control over the environmental stress  
25 responses of a plant can be achieved by the present invention.

In each of the above embodiments, the binding protein comprises an  
amino acid sequence which is capable of binding to a DNA regulatory  
sequence which regulates one or more environmental stress tolerance genes.  
In a preferred embodiment, the binding protein further comprises a transcription  
30 activation region which acts in concert with the binding sequence to regulate  
expression of one or more environmental stress tolerance genes in the plant by  
regulating expression of one or more environmental stress tolerance genes.  
The environmental stress tolerance gene, DNA regulatory sequence, sequence  
encoding the binding sequence, and the sequence encoding the transcription

activation region may each independently be native or non-native to the plant and may each independently be homologous or heterologous relative to each other.

Optionally, the binding protein satisfies one or more of the following requirements:

the binding protein comprises an AP2 domain which comprises a consensus sequence sufficiently homologous to any one of the consensus sequences shown in Figures 19A, 19B, or 19C that the binding protein is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises an AP2 domain which comprises a consensus sequence shown in Figures 19A, 19B or 19C;

the binding protein comprises an AP2 domain which comprises the amino acid residues shown in Figures 19D or 19E;

the binding protein comprises an AP2 domain which is sufficiently homologous to at least one of the AP2 domains shown in the application such that it is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises one of the AP2 domain sequences shown in this application, including, but not limited to SEQ. I.D. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95;

the binding protein comprises a sequence which comprises one of the amino terminus domains shown in Figure 20 (it is noted that the sequence need not be at the amino terminus of the binding protein);

the binding protein comprises the consensus sequence for the amino terminus domains shown in Figure 20, (it is noted that the sequence need not be at the amino terminus of the binding protein);

the binding protein comprises a sequence which comprises one of the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein); and



the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21B (it is noted that the sequence need not be at the carboxy terminus of the binding protein).

5 The amino acid sequence encoding the binding protein may be a naturally occurring sequence such as the ones shown in SEQ. ID. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95 or may be a non-naturally occurring sequence. It is noted, however, that binding proteins according to the present invention are intended to encompass non-naturally occurring sequences which  
10 are derivatives of the classes of binding proteins taught herein. For example, additional binding proteins may be constructed using one of the AP2 domains taught herein or the consensus sequence of these AP2 domains. It may be desirable to include with the AP2 domain a transcription activation region. The transcription activation region may be native to the plant or non-native to the  
15 plant in which the binding protein will be used. For example, the sequence may include a subsequence which encodes a binding domain for the DNA regulatory sequence fused to a transcription activating region, such as the transcription activating region of VP16 or GAL4. Optionally, one can include in the binding protein one of the amino terminus domains, the consensus  
20 sequence for the amino terminus domain, one of the carboxy terminus domains and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

25 Optionally, the binding protein can be viewed as comprising one of the amino terminus domains, the consensus sequence for the amino terminus domain, one of the carboxy terminus domains and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein  
30 and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

A method is also provided for identifying from a cDNA library of at least a portion of a plant genome a gene sequence encoding a protein capable of

binding to a target DNA regulatory sequence. In one embodiment, the method comprises

- 5 taking a microorganism which includes a target DNA regulatory sequence for one or more environmental stress tolerance genes, a transcription activator for activating expression of a reporter gene, and a reporter gene whose expression is activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and an activation domain capable of activating the transcription activator;
- 10 fusing sequences from a cDNA library of at least a portion of a plant genome to a sequence which encodes a functional activation domain in the microorganism;
- introducing the fused sequences into the microorganism; and
- selecting microorganisms which express the reporter gene, expression of the reporter gene indicating expression of a fusion protein which includes a binding domain for the target DNA regulatory sequence and the activation domain; and
- 15 identifying the gene sequence from the cDNA library introduced into the microorganism.

The target DNA regulatory sequence may optionally include the subsequence CCG or the subsequence CCGAC. This embodiment of the invention also relates to DNA in substantially isolated form, nucleic acid constructs capable of transforming a plant, cells, and transformed plants which include a gene sequence identified by this method.

While the present invention is described with regard to the use of binding proteins which can bind to a DNA regulatory sequence that regulates environmental stress tolerance genes in a plant, it is noted that these same binding proteins can also be used to regulate genes other than environmental stress tolerance genes by placing these other genes under the regulatory control of the DNA regulatory sequence. For example, protein kinases that induce cold and drought inducible genes can be regulated by placing a protein kinase gene under the control of a promoter whose expression is regulated by the DNA regulatory sequence. PCT/US97/23019 (Intl Publication Number WO 98/26045) describes protein kinases that when constitutively expressed, induce cold and drought inducible genes. The ATCDPK1a and the ATCDPK1

constitutive protein kinase coding regions (PCT/US97/23019) can be isolated by PCR and inserted into the drought and cold inducible promoters described in Example 8 by one skilled in the art. The expression of these ATCDPK1 constitutive protein kinase coding regions (PCT/US97/23019) from the drought and cold inducible promoters will increase the drought and cold tolerance of plants and should be synergistic with the drought and cold tolerance induced by CBF expression under inducible promoters.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B show how the yeast reporter strains were constructed.

Figure 1A is a schematic diagram showing the screening strategy.

Figure 1B is a chart showing activity of the "positive" cDNA clones in yeast reporter strains.

Figures 2A, 2B, 2C and 2D provide an analysis of the pACT-11 cDNA clone.

Figure 2A is a schematic drawing of the pACT-11 cDNA insert indicating the location and 5' to 3' orientation of the 24 kDa polypeptide and 25S rRNA sequences.

Figure 2B is a DNA and amino acid sequence of the 24 kDa polypeptide (SEQ ID NO:1 and SEQ ID NO:2).

Figure 2C is a schematic drawing indicating the relative positions of the potential nuclear localization signal (NLS), the AP2 domain and the acidic region of the 24 kDa polypeptide.

Figure 2D is a chart showing comparison of the AP2 domain of the 24 kDa polypeptide with that of the tobacco DNA binding protein EREBP2.

Figure 3 is a chart showing activation of reporter genes by the 24 kDa polypeptide.

Figure 4 is a photograph of an electrophoresis gel showing expression of the recombinant 24 kDa polypeptide in *E. coli*.

Figure 5 is a photograph of a gel for shift assays indicating that CBF1 binds to the C-repeat/DRE.

Figure 6 is a photograph of a southern blot analysis indicating *CBF1* is a unique or low copy number gene.

Figures 7A, 7B and 7C relate to *CBF1* transcripts in control and cold-treated *Arabidopsis*.

5        Figure 7A is a photograph of a membrane RNA isolated from *Arabidopsis* plants that were grown at 22 C or grown at 22 C and transferred to 2.5 C for the indicated times.

Figure 7B is a graph showing relative transcript levels of *CBF1* in control and cold-treated plants.

10       Figure 7C is a graph showing relative transcript levels of *COR15a* in control and cold-treated plants.

Figure 8 is a Northern blot showing *CBF1* and *COR* transcript levels in RLD and transgenic *Arabidopsis* plants.

15       Figure 9 is an immunoblot showing *COR15a* protein levels in RLD and transgenic *Arabidopsis* plants.

Figures 10A and 10B are graphs showing freezing tolerance of leaves from RLD and transgenic *Arabidopsis* plants.

Figure 11 is a photograph showing freezing survival of RLD and A6 *Arabidopsis* plants.

20       Figure 12 shows the DNA sequence for *CBF2* encoding CBF2.

Figure 13 shows the DNA sequence for *CBF3* encoding CBF3.

Figure 14 shows the amino acid alignment of proteins CBF1, CBF2 and CBF3.

25       Figure 15 is a graph showing transcription regulation of *COR* genes by *CBF1*, *CBF2* and *CBF3* genes in yeast.

Figure 16 shows the amino acid sequence of a canola homolog and its alignment to the amino acid sequence of CBF1.

30       Figures 17A, 17B, 17C, 17D, 17E, 17F and 17G show restriction maps of plasmids pMB12008, pMB12009, pMB12010, pMB12011, pMB12012, pMB12013, and pMB12014, respectively.

Figure 18A shows the DNA sequences for the CBF homologs from *Brassica juncea*, *Brassica napus*, *Brassica oleracea*, *Brassica rapa*, *Glycine max*, *Raphanus sativus* and *Zea Maize*.

Figure 18B shows the amino acid sequences (one-letter abbreviations) encoded by the DNA sequences (shown in FIG. 18A) for CBF homologs from *Brassica juncea*, *Brassica napus*, *Brassica oleracea*, *Brassica rapa*, *Glycine max*, *Raphanus sativus* and *Zea Maize*.

5        Figure 19A shows an amino acid alignment of the AP2 domains of several CBF proteins with the consensus sequence between the proteins highlighted as well as a comparison of the AP2 domains with that of the tobacco DNA binding protein EREBp2.

10        Figure 19B shows an amino acid alignment of the AP2 domains of several CBF proteins including *dreb2a* and *dreb2b* with the consensus sequence between the proteins highlighted.

Figure 19C shows an amino acid alignment of the AP2 domains of several CBF proteins including *dreb2a*, *dreb2b*, and *tiny* with the consensus sequence between the proteins highlighted.

15        Figure 19D shows a difference between the consensus sequence shown in Figure 19A and *tiny*.

Figure 19E shows a difference between the consensus sequence shown in Figure 19B and *tiny*.

20        Figure 20 shows an amino acid alignment of the amino terminus of several CBF proteins with their consensus sequence highlighted.

Figure 21A and 21B show an amino acid alignment of the carboxy terminus of several CBF proteins, with their consensus sequences highlighted.

## 25        DETAILED DESCRIPTION

25        The present invention relates to DNA encoding binding proteins capable of binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The present invention also relates to the binding proteins encoded by the DNA. The DNA and binding  
30        proteins may be native or non-native relative to the DNA regulatory sequence of the plant. The DNA and binding proteins may also be native or non-native relative to environmental stress tolerance genes of the plant which are regulated by the DNA regulatory sequence.

The present invention also relates to methods for using the DNA and binding proteins to regulate expression of one or more native or non-native environmental stress tolerance genes in a plant. These methods may include introducing DNA encoding a binding protein capable of binding to a DNA regulatory sequence into a plant, introducing a promoter into a plant which regulates expression of the binding protein, introducing a DNA regulatory sequence into a plant to which a binding protein can bind, and/or introducing one or more environmental stress tolerance genes into a plant whose expression is regulated by a DNA regulatory sequence.

The present invention also relates to recombinant cells, plants and plant materials (e.g., plant tissue, seeds) into which one or more gene sequences encoding a binding protein have been introduced as well as cells, plants and plant materials within which recombinant binding proteins encoded by these gene sequences are expressed. By introducing a gene sequence encoding a binding protein into a plant, a binding protein can be expressed within the plant which regulates expression of one or more stress tolerance genes in the plant. Regulation of expression can include causing one or more stress tolerance genes to be expressed under different conditions than those genes would be in the plant's native state, increasing a level of expression of one or more stress tolerance genes, and/or causing the expression of one or more stress tolerance genes to be inducible by an exogenous agent. Expression of the binding protein can be under the control of a variety of promoters. For example, promoters can be used to overexpress the binding protein, change the environment conditions under which the binding protein is expressed, or enable the expression of the binding protein to be induced, for example by the addition of an exogenous inducing agent. Promoters can also be used to cause the protein to be expressed at selected times during a plant's life. Tissue-specific promoters can be used to cause the protein to be expressed in selected tissues. For example, flower-, fruit- and seed-specific promoters can be used to cause the protein to be selectively expressed in flowers, fruits or seeds of the plant.

The present invention also relates to cells, recombinant plants and plant materials into which a recombinant promoter is introduced which controls a level of expression of one or more gene sequences encoding a binding protein.

The one or more gene sequences may be recombinant native or non-native sequences or may be native, non-recombinant gene sequences whose expression is altered by the introduction of the recombinant promoter.

5 The present invention also relates to cells, recombinant plants and plant materials into which a recombinant native or non-native DNA regulatory sequence is introduced which regulates expression of one or more native or non-native environmental stress tolerance genes.

10 Examples of environmental stresses for which stress tolerance genes are known to exist include, but are not limited to, cold tolerance, dehydration tolerance, and salinity tolerance. As used herein, environmental stress tolerance genes refer to genes which function to acclimate a plant to an environment stress. For example, cold tolerance genes, also referred to as COR genes (COld Regulated), refer to genes which function to acclimate a plant to a cold temperature environment. These genes typically are activated  
15 when a plant is exposed to cold temperatures. Dehydration tolerance genes refer to genes which function to acclimate a plant to dehydration stress. These genes typically are activated in response to dehydration conditions which can be associated with drought or cold temperatures which cause water in the plant to freeze and thereby dehydrate the plant tissue. It is noted that some cold  
20 tolerance genes may function to provide a plant with a degree of dehydration tolerance and visa versa. For example, COR genes are known to also be activated by dehydration stress. This application is intended to encompass genes which regulate one or more environmental stress tolerance genes such as cold tolerance genes, dehydration tolerance genes, and genes which  
25 perform a dual function of cold and dehydration tolerance.

One embodiment of the invention relates to a DNA sequence in isolated form which includes a sequence encoding a binding protein capable of selectively binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. The binding  
30 protein is preferably capable of regulating expression of one or more environmental stress tolerance genes in a plant by selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. In one variation, the binding protein is a non-naturally occurring protein formed by combining an amino acid sequence capable of

binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence with an amino acid sequence which forms a transcription activation region which regulates expression of one or more environmental stress tolerance genes in a plant by regulating expression of one or more  
5 environmental stress tolerance genes when the binding protein binds to the regulatory region.

The DNA sequence may exist in a variety of forms including a plasmid or vector and can include sequences unrelated to the gene sequence encoding the binding protein. For example, the DNA sequence can include a promoter  
10 which regulates expression of the regulatory gene.

In one variation of this embodiment, the DNA regulatory sequence is a C-repeat cold and drought regulation element (C-repeat/DRE). As will be explained and demonstrated herein, C-repeat/DRE regulatory sequences appear to be conserved in plants with some degree of variability plant to plant.  
15 Using the teachings of the present invention, C-repeat/DRE regulatory sequences native to different plants can be identified as well as the native stress tolerance regulatory genes which encode for proteins which bind to the C-repeat/DRE DNA regulatory sequences. Hence, although the examples provided herein to describe the present invention are described with regard to  
20 the *Arabidopsis* C-repeat/DRE DNA regulatory sequence, the present invention is not intended to be limited to the *Arabidopsis* C-repeat/DRE DNA regulatory sequence. Rather, the *Arabidopsis* C-repeat/DRE DNA regulatory sequence is believed to be a member of a class of environmental stress response regulatory elements which includes the subsequence CCGAC which  
25 in turn is believed to be a member of a class of environmental stress response regulatory elements which includes the subsequence CCG. Other different classes of environmental stress response regulatory elements may also exist. The teachings of the present invention may be used to identify sequences which bind to these and other classes of environmental stress response  
30 regulatory elements once they are identified.

In one variation of this embodiment, the gene sequence encodes a binding protein which selectively binds to a member of a class of DNA regulatory sequences which includes the subsequence CCG. In another variation, the gene sequence encodes a binding protein which selectively binds



to a member of a class of DNA regulatory sequences which includes the subsequence CCGAC. The CCGAC subsequence has been found to present in the C-repeat/DRE DNA regulatory sequences of *Arabidopsis* and *Brassica* and to function in Tobacco based on the ability of the C-repeat/DRE to direct cold and tolerance regulated gene expression.

In yet another variation, the stress tolerance regulatory gene sequence encodes a binding protein which includes an AP2 domain. It is believed that a significant class of environmental stress tolerance regulatory genes encode for binding proteins with an AP2 domain capable of binding to the DNA regulatory sequence. The AP2 domain of the binding protein is preferably a homolog of the AP2 domain of one of the CBF binding proteins described herein. The subsequence encoding the AP2 domain is preferably a homolog of a subsequence of one of the CBF genes described herein which encodes an AP2 domain.

In another variation, the DNA sequence encoding the binding protein satisfies one or more of the following requirements:

the binding protein comprises an AP2 domain which comprises a consensus sequence sufficiently homologous to any one of the consensus sequences shown in Figures 19A, 19B, or 19C that the binding protein is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises an AP2 domain which comprises a consensus sequence shown in Figures 19A, 19B or 19C;

the binding protein comprises an AP2 domain which comprises the amino acid residues shown in Figures 19D or 19E;

the binding protein comprises an AP2 domain which is sufficiently homologous to at least one of the AP2 domains shown in the application such that it is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises one of the AP2 domain sequences shown in this application, including, but not limited to SEQ. I.D. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95;

the binding protein comprises a sequence which comprises one of the amino terminus domains shown in Figure 20 (it is noted that the sequence need not be at the amino terminus of the binding protein);

5 the binding protein comprises the consensus sequence for the amino terminus domains shown in Figure 20, (it is noted that the sequence need not be at the amino terminus of the binding protein);

the binding protein comprises a sequence which comprises one of the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

10 the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21B (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

15 one of SEQ. I.D. Nos. 1, 12, 14, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, and 94; or

a sequence which has substantially the same degree of homology to SEQ. I.D. Nos. 1, 12, 14, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 20 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, and 94 as these sequences have with each other.

The present invention also relates to a method for identifying gene sequences from at least a portion of a plant genome which encode binding proteins capable of binding to a target DNA regulatory sequence which 25 regulates expression of one or more stress tolerance genes in a plant.

In one embodiment, the method includes:

taking a microorganism which includes a target DNA regulatory sequence for one or more environmental stress tolerance genes, a transcription activator for activating expression of a reporter gene, and a 30 reporter gene whose expression is activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and an activation domain capable of activating the transcription activator;

fusing sequences from a cDNA library of at least a portion of a plant genome to a sequence which encodes a functional activation domain in the microorganism;

introducing the fused sequences into the microorganism; and

5 selecting microorganisms which express the reporter gene, expression of the reporter gene indicating expression of a fusion protein which includes a binding domain for the target DNA regulatory sequence and the activation domain; and

10 identifying the gene sequence from the cDNA library introduced into the microorganism.

In one variation of the method, the target DNA regulatory sequence includes the subsequence CCG and in another embodiment includes the subsequence CCGAC. In yet another variation, the target DNA regulatory sequence is the C-repeat/DRE for *Arabidopsis*. According to the above  
15 method, the target DNA regulatory sequence is preferably native to the plant family and more preferably to the plant species from which the cDNA library is derived.

In another variation of this embodiment, the cDNA library used in the method consists of sequences which encode for a protein having an AP2  
20 domain since it is believed that a significant class of genes encoding binding proteins for stress tolerance genes encode an AP2 domain. As will be explained herein, screening for DNA sequences from a plant genome which exhibit this functional feature has been shown to be effective for isolating gene sequences encoding binding proteins of the present invention.

25 In another variation of this method, the sequences from the cDNA library are fused to a sequence which includes a selectable marker, the method further including the step of selecting for microorganisms expressing the selectable marker.

30 While the above methodology of the present invention is described herein with regard to identifying binding protein gene sequences from *Arabidopsis* cDNA using the C-repeat/DRE regulatory sequence for *Arabidopsis*, it is noted that this methodology can be readily used to identify regulatory binding protein gene sequences for other plants by using a DNA regulatory sequence native to those plants. Alternatively, different

permutations of the CCG subsequence can be used as the target DNA regulatory sequence.

5 An example of a microorganism which may be used in the above method is yeast. cDNA can be introduced into the microorganism by a variety of mechanisms including plasmids and vectors. In one particular embodiment, the reporter gene is beta-galactosidase.

The present invention also relates to any DNA sequences and binding proteins encoded by those DNA sequences which are identified by the above screening method.

10 The present invention also relates to a protein expressed by an environmental stress tolerance regulatory gene according to the present invention which can function *in vivo* in a plant to regulate expression of one or more environmental stress tolerance genes.

15 According to one embodiment, the protein is a recombinant binding protein expressed by a copy of a recombinant gene which is either not native to the plant or is native to the plant but introduced into the plant by recombinant methodology. For example, one might wish to introduce one or more copies of a regulatory gene which is native to the plant but is under the control of a promoter which overexpresses the binding protein, expresses the binding protein independent of an environmental stress, expresses the binding protein at a higher level in response to the same environmental stress than would a plant in its native state, expresses the binding protein in response to different environmental stress conditions, and/or be induced to express the binding protein by an exogenous agent to which the plant can be exposed.

20 Alternatively, one might wish to introduce one or more copies of a regulatory gene which is not native to the plant. For example, the non-native regulatory gene may be used to alter the way in which native environmental stress tolerance genes are regulated. Alternatively, the non-native regulatory gene may be used to regulate environmental stress tolerance genes which are also not native to the plant. The non-native regulatory gene may be used to bind to a DNA regulatory region which is not native to the plant.

25 In another embodiment, the proteins have been isolated from a recombinant organism. The organism may be a microorganism (e.g., bacteria,

yeast) or a multicellular organism such as a plant. In one variation, the protein is in substantially isolated form.

5 In yet another embodiment, the protein is a native, non-recombinant binding protein whose expression is regulated within a plant by a recombinant native or non-native promoter. For example, one might wish to replace a native promoter with a recombinant promoter which overexpresses the binding protein, expresses the binding protein independent of an environmental stress, expresses the binding protein at a higher level in response to the same environmental stress than would a plant in its native state, expresses the binding protein in response to different environmental stress conditions, and/or  
10 be induced to express the binding protein by an exogenous agent to which the plant can be exposed.

In one variation of the above embodiments, the protein is capable of selectively binding to a DNA regulatory sequence for one or more  
15 environmental stress tolerance genes in a plant. In another variation, the protein includes an AP2 domain which is capable of selectively binding to a DNA regulatory sequence for one or more environmental stress tolerance genes in a plant. One method which may be used to determine whether the protein binds selectively to the DNA regulatory sequence is a gel shift assay.  
20 The DNA regulatory sequence may optionally include a CCG subsequence, a CCGAC subsequence and optionally the C-repeat / DRE sequence of *Arabidopsis*.

In another variation of the above embodiments, the binding protein satisfies one or more of the following requirements:

25 the binding protein comprises an AP2 domain which comprises a consensus sequence sufficiently homologous to any one of the consensus sequences shown in Figures 19A, 19B, or 19C that the binding protein is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

30 the binding protein comprises an AP2 domain which comprises a consensus sequence shown in Figures 19A, 19B or 19C;

the binding protein comprises an AP2 domain which comprises the amino acid residues shown in Figures 19D or 19E;

the binding protein comprises an AP2 domain which is sufficiently homologous to at least one of the AP2 domains shown in the application such that it is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

5 the binding protein comprises one of the AP2 domain sequences shown in this application, including, but not limited to SEQ. I.D. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95;

10 the binding protein comprises a sequence which comprises one of the amino terminus domains shown in Figure 20 (it is noted that the sequence need not be at the amino terminus of the binding protein);

the binding protein comprises the consensus sequence for the amino terminus domains shown in Figure 20, (it is noted that the sequence need not be at the amino terminus of the binding protein);

15 the binding protein comprises a sequence which comprises one of the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein); and

20 the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21B (it is noted that the sequence need not be at the carboxy terminus of the binding protein).

25 The sequence of the binding protein may be a naturally occurring sequence such as the ones shown in SEQ. ID. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95 or may be a non-naturally occurring sequence. It is noted, however, that binding proteins according to the present invention are intended to encompass non-naturally occurring sequences which are derivatives of the

30 classes of binding proteins taught herein. For example, additional binding proteins may be constructed using one of the AP2 domains taught herein or the consensus sequence of these AP2 domains. It may be desirable to include with the AP2 domain a transcription activation region. The transcription activation region may be native to the plant or non-native to the plant in which

the binding protein will be used. For example, the sequence may include a subsequence which encodes a binding domain for the DNA regulatory sequence fused to a transcription activating region, such as the transcription activating region of VP16 or GAL4. Optionally, one can include in the binding protein one of the amino terminus domains, the consensus sequence for the amino terminus domain, one of the carboxy terminus domains and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

Optionally, the binding protein can be viewed as comprising one of the amino terminus domains, the consensus sequence for the amino terminus domain, one of the carboxy terminus domains and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

In another embodiment, the binding protein is an isolated protein or a recombinantly produced protein which has a molecular weight of about 26 kDa as measured in an electrophoresis gel and binds to a DNA regulatory sequence which regulates a cold or dehydration regulated gene of *Arabidopsis thaliana*.

The present invention also relates to DNA and RNA constructs, such as plasmids, vectors, and the like, which are capable of transforming a plant. The constructs include a sequence which encodes a binding protein capable of selectively binding to a DNA regulatory sequence which regulates the one or more environmental stress tolerance genes. The binding protein is preferably able to regulate expression of one or more environmental stress tolerance genes in a plant by selectively binding to the DNA regulatory sequence. More preferably, when transformed into a plant, the sequence regulates expression of one or more environmental stress tolerance genes in the plant by expressing the binding protein. In one embodiment, the DNA construct includes a promoter and a regulatory gene sequence whose expression is under the control of the promoter. Different promoters may be used to select the degree of expression

or conditions under which the regulatory gene is expressed. For example, the promoter can be used to cause overexpression of the regulatory gene, expression of the regulatory gene independent of an environmental stress, expression of the regulatory gene at a higher level in response to the same  
5 environmental stress than would a plant in its native state, expression of the regulatory gene in response to different environmental stress conditions, and/or induction of expression of the regulatory gene by an exogenous agent to which the plant can be exposed.

Promoters can also be used to cause the protein to be expressed at selected  
10 times during a plant's life. Tissue-specific promoters can be used to cause the protein to be expressed in selected tissues. For example, flower-, fruit- and seed-specific promoters can be used to cause the protein to be selectively expressed in flowers, fruits or seeds of the plant.

In another embodiment, the DNA construct comprises a sequence  
15 which encodes:

a binding protein comprising an AP2 domain which comprises a consensus sequence sufficiently homologous to any one of the consensus sequences shown in Figures 19A, 19B, or 19C that the binding protein is capable of binding to a CCG regulatory sequence, preferably a CCGAC  
20 regulatory sequence;

a binding protein comprising an AP2 domain which comprises a consensus sequence shown in Figures 19A, 19B or 19C;

a binding protein comprising an AP2 domain which comprises the amino acid residues shown in Figures 19D or 19E;

25 a binding protein comprising an AP2 domain which is sufficiently homologous to at least one of the AP2 domains shown in the application such that it is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

30 a binding protein comprising one of the AP2 domain sequences shown in this application, including, but not limited to SEQ. I.D. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95;



a binding protein comprising a sequence which comprises one of the amino terminus domains shown in Figure 20 (it is noted that the sequence need not be at the amino terminus of the binding protein);

5 a binding protein comprising the consensus sequence for the amino terminus domains shown in Figure 20, (it is noted that the sequence need not be at the amino terminus of the binding protein);

a binding protein comprising a sequence which comprises one of the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

10 a binding protein comprising the consensus sequence for the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

a binding protein comprising the consensus sequence for the carboxy terminus domains shown in Figure 21B (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

15 one of SEQ. I.D. Nos. 1, 12, 14, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, and 94; or

a sequence which has substantially the same degree of homology to SEQ. I.D. Nos. 1, 12, 14, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 20 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, and 94 as these sequences have with each other.

The present invention also relates to plasmids pCBF1 (ATCC 98063), pCBF2, and pCBF3.

25 The present invention also relates to a recombinant microorganism, such as a bacterium, yeast, fungus, virus, into which at least one copy of a regulatory gene encoding a binding protein of the present invention has been introduced by a recombinant methodology.

30 The present invention also relates to recombinant plants into which at least one copy of a regulatory gene encoding a binding protein of the present invention has been introduced by a recombinant methodology. The recombinant copy of the regulatory gene may be native or non-native to the plant and express a binding protein which is either native or non-native to the plant. Expression of the recombinant copy of the regulatory gene may be under the control of the promoter. The promoter may increase the level at

which the regulatory gene is expressed, express the regulatory gene without being induced by an environmental stress and/or express the regulatory gene in response to a different form or degree of environmental stress than would otherwise be needed to induce expression of the regulatory gene. For  
5 example, a promoter can be used which turns on at a temperature that is warmer than the temperature at which the plant normally exhibits cold tolerance. This would enable the cold tolerance thermostat of a plant to be altered. Similarly, a promoter can be used which turns on at a dehydration condition that is wetter than the dehydration condition at which the plant  
10 normally exhibits dehydration tolerance. This would enable the level at which a plant responds to dehydration to be altered. A promoter can also be used which causes a higher level of expression to occur at a given environmental condition (e.g. temperature and/or dryness) than the plant would express in its native state. The promoter may also be inducible by an exogenous agent, i.e.,  
15 express the regulatory gene in response to the presence of an agent to which the promoter is exposed. This would enable stress tolerance to be induced by applying an inducing agent to the plant.

Selection of the promoter can also be used to determine what tissues in the plant express the binding protein as well as when expression occurs in the  
20 plant's lifecycle. By selecting a promoter which regulates in what tissues and when in a plant's life the promoter functions to regulate expression of the binding protein, in combination with the selecting how that promoter regulates expression (level of expression and/or type of environmental or chemical induction), an incredible range of control over the environmental stress  
25 responses of a plant can be achieved according to the present invention. For example, flower-, fruit- and seed-specific promoters can be used to regulate the expression of the binding protein in these tissues of the plant, especially when sudden frosts strike in early spring and late fall.

The environmental stress tolerance gene regulated by the  
30 recombinantly expressed regulatory gene may be native or non-native to the plant. Hence, in one embodiment, the plant includes a recombinant copy of a regulatory gene which is native to the plant and expresses a native protein which functions within the plant to regulate expression of a native environmental stress tolerance gene. In this embodiment, the recombinant

plant expresses a higher level of the native regulatory gene than the plant would otherwise.

5 In another embodiment, at least one of the regulatory genes and the environmental stress tolerance genes is not native to the plant. For example, the regulatory gene can be native and the environmental stress tolerance gene is non-native, or the regulatory gene is non-native and the environmental stress tolerance gene is native to the plant.

10 In yet another embodiment, the plant can include a recombinant copy of a regulatory gene which is not native to the plant as well as a recombinant copy of one or more environmental stress tolerance genes which also is not native to the plant. According to this embodiment, the non-native regulatory gene expresses a non-native binding protein which functions within the plant to regulate expression of the one or more non-native environmental stress tolerance genes. In this regard, it is envisioned that the present invention can be used to introduce, change and/or augment the environmental stress tolerance of a plant by introducing and causing the expression of environmental stress tolerance which the plant does not have in its native form. Accordingly, plants from warmer climates can be engineered to include one or more cold tolerance genes along with a regulatory gene needed to cause expression of the cold tolerance genes in the plant so that the engineered plant can survive better in a colder climate. Similarly, a plant can be engineered to include one or more dehydration tolerance genes along with a regulatory gene needed to cause expression of the dehydration tolerance gene so that the engineered plant can grow better in a dryer climate. In this regard, it should be possible to take a plant which grows well in a first climate and engineer it to include stress tolerance genes and regulatory genes native to a second climate so that the plant can grow well in the second climate.

The present invention also relates to a method for changing or enhancing the environmental stress tolerance of a plant.

30 In one embodiment, the method includes introducing at least one copy of a regulatory gene encoding a binding protein of the present invention into a plant; expressing the binding protein encoded by the regulatory gene; and using the expressed binding protein to stimulate expression of at least one environmental stress tolerance gene through binding to a DNA regulatory

sequence. According to this embodiment, the regulatory gene may be non-recombinant or recombinant native or non-native to the plant. Similarly, the DNA regulatory sequence and the environmental stress tolerance gene may each independently be native or non-native to the plant. In one variation of this  
5 embodiment, the method further includes recombinantly introducing an environmental stress tolerance gene into the plant which is regulated by the recombinant regulatory gene.

In another embodiment, the method includes introducing a recombinant promoter which regulates expression of a regulatory gene encoding a binding  
10 protein of the present invention into a plant; expressing the binding protein under the control of the recombinant promoter in the plant; and using the expressed binding protein to stimulate expression of at least one environmental stress tolerance gene through binding to a DNA regulatory sequence.

According to this embodiment, the regulatory gene, the DNA regulatory  
15 sequence and the environmental stress tolerance gene may each independently be non-recombinant or recombinant native or non-native to the plant. Also according to this embodiment, the promoter can be a tissue-specific promoter such as a flower-, fruit- and seed-specific promoter. In this instance, expressing the binding protein includes selectively expressing the  
20 binding protein in a particular type of tissue, such as flowers, fruits or seeds of the plant.

In yet another embodiment, the method includes introducing at least one recombinant environmental stress tolerance gene into a plant; expressing a binding protein; and using the expressed binding protein to stimulate  
25 expression of the recombinant environmental stress tolerance gene through binding to a DNA regulatory sequence. According to this embodiment, the gene encoding the regulatory protein, and the DNA regulatory sequence may each independently be non-recombinant or recombinant native or non-native to the plant. The recombinant environmental stress tolerance gene may be either  
30 native or non-native to the plant.

## 1. DEFINITIONS

5 The term "C-repeat cold and drought regulation element" or "C-repeat/DRE" refers to a sequence which includes CCG and functions as a binding domain in a plant to regulate expression of one or more environmental stress tolerance genes, such as cold or dehydration stress tolerance genes.

The term "cold stress" refers to a decrease in ambient temperature, including a decrease to freezing temperatures, which causes a plant to attempt to acclimate itself to the decreased ambient temperature.

10 The term "dehydration stress" refers to drought, high salinity and other conditions which cause a decrease in cellular water potential in a plant.

Transformation means the process for changing the genotype of a recipient organism by the stable introduction of DNA by whatever means.

15 A transgenic plant is a plant containing DNA sequences which were introduced by transformation. Horticultural and crop plants particularly benefit from the present invention.

Translation means the process whereby the genetic information in an mRNA molecule directs the order of specific amino acids during protein synthesis.

20 The term "essentially homologous" means that the DNA or protein is sufficiently duplicative of that set forth in Figure 2B to produce the same result. Such DNA can be used as a probe to isolate DNA's in other plants.

25 A promoter is a DNA fragment which causes transcription of genetic material. For the purposes described herein, promoter is used to denote DNA fragments that permit transcription in plant cells.

A poly-A addition site is a nucleotide sequence which causes certain enzymes to cleave mRNA at a specific site and to add a sequence of adenylic acid residues to the 3'-end of the mRNA.

30 The phrase "DNA in isolated form" refers to DNA sequence which has been at least partially separated from other DNA present in its native state in an organism. A cDNA library of genomic DNA is not "DNA in isolated form" whereas DNA which has been at least partially purified by gel electrophoresis corresponds to "DNA in isolated form".

## 2. C-Repeat/DRE Regulatory Elements In Plants

C-repeat cold and drought regulation elements (C-repeat/DRE) are sequences which function as a *cis*-acting regulatory element that stimulates transcription in response to an environmental stress, such as low temperature (Yamaguchi-Shinozaki, K., et al., *The Plant Cell* 6:251-264 (1994); and Baker, S. S., et al., *Plant Mol. Biol.* 24:701-713 (1994); Jiang, C., et al., *Plant Mol. Biol.* 30:679-684 (1996)) or dehydration stress and high salinity (Yamaguchi-Shinozaki, K., et al., *The Plant Cell* 6:251-264 (1994)). An object of the research leading to the present invention was the determination of how a C-repeat/DRE stimulates gene expression in response to these environmental factors, and whether cold, dehydration and high salinity affect independent or overlapping regulatory systems.

The first step toward determining how a C-repeat/DRE regulation element stimulates gene expression was the identification of the C-repeat cold and drought regulation element itself. The 5 base pair core sequence, CCGAC, has been found to be present once to multiple times in a variety of plant cold-regulated promoters in *Arabidopsis* and *Brassica* including the *COR15a* (Baker, S. S., et al, *Plant. Mol. Biol.* 24:701-713 (1994)); *COR78/RD29A* (Horvath, D. P., et al, *Plant Physiol.* 103:1047-1053 (1993) and Yamaguchi-Shinozaki, K., et al., *The Plant Cell* 6:251-264 (1994)); *COR6.6* (Wang, H., et al., *Plant Mol. Biol.* 28:605-617 (1995)); and *KIN1* (Wang, H., et al, *Plant Mol. Biol.* 28:605-617 (1995)) genes of *Arabidopsis* and the *BN115* gene of *Brassica napus* (White, T. C., et al, *Plant Physiol.* 106:917-928 (1994)). As shown in the examples herein, core sequence CCGAC was used to identify proteins encoded by genes within the *Arabidopsis* genome which bind to this core sequence.

Applicants believe that the CCGAC core sequence is a member of family of core sequences having the common subsequence CCG. The binding of CBF1 to the C-repeat/DRE involves the AP2 domain. In this regard, it is germane to note that the tobacco ethylene response element, AGCCGCC, closely resembles the C-repeat/DRE sequences present in the promoters of the *Arabidopsis* genes *COR15a*, GGCCGAC, and *COR78/RD29A*, TACCGAC. While the specific teachings in the present invention used only a DNA

regulatory sequence which includes a CCGAC subsequence as the C-repeat/DRE core regulatory sequence, Applicants believe that other C-repeat/DRE regulatory sequences exist which belong to a broader CCG family of regulatory sequences. By screening plant genomes according to the methodology taught herein using other members of the CCG family, additional regulatory sequences as well as the binding proteins which bind to these regulatory sequences can be identified. For example, plants which are known to exhibit a form of environmental stress tolerance can be screened according to the blue colony assay and other screening methodologies used in the present invention with other members of the CCG family in order to identify other binding proteins and their gene sequences. Examples of other members of the CCG family include, but are not limited to, environmental stress response regulatory elements which include one of the following sequences: CCGAA, CCGAT, CCGAC, CCGAG, CCGTA, CCGTT, CCGTC, CCGTG, CCGCA, CCGCT, CCGCG, CCGCC, CCGGA, CCGGT, CCGGC, CCGGG, AACCG, ATCCG, ACCCG, AGCCG, TACCG, TTCCG, TCCCG, TGCCG, CACCG, CTCCG, CGCCG, CCCC G, GACCG, GTCCG, GCCCG, GGCCG, ACCGA, ACCGT, ACCGC, ACCGG, TCCGA, TCCGT, TCCGC, TCCGG, CCCGA, CCCGT, CCCGC, CCCGG, GCCGA, GCCGT, GCCGC, and GCCGG.

Applicants also believe that other families of environmental stress tolerance DNA regulatory sequences, other than the CCG family may exist. The methodologies of the present invention may be used once such other families are identified in order to identify specific environmental stress tolerance DNA regulatory sequences and associated binding proteins.

### **3. Identification Of Environmental Stress Tolerance Regulatory Gene Sequences Using Target Regulatory Sequence**

It is possible to take a cDNA library of at least a portion of a plant genome and screen the cDNA library for the presence of regulatory gene sequences which encode binding proteins capable of binding to a target regulatory sequence. As used here, a target DNA regulatory sequence refers to a sequence to which a binding protein for one or more environmental stress tolerance genes binds. Permutations of the CCG and CCGAC families of DNA regulatory sequences represent examples of target DNA regulatory sequences.

As detailed in Example 1 herein, this was the approach was used to identify CBF1, a sequence which encodes a binding protein for the *Arabidopsis* DNA regulatory sequence, from an *Arabidopsis* cDNA library .

5 First a target regulatory sequence is selected. The target regulatory sequence is preferably native to the plant from which the cDNA library being screened is derived.

Once a target regulatory sequence is selected, the target regulatory sequence is fused to a reporter gene and introduced into a microorganism. Expression of the reporter gene can be activated by a protein which includes a binding domain capable of binding to the target DNA regulatory sequence and  
10 an activation domain capable of activating transcription.

Sequences from a cDNA library of at least a portion of a plant genome are then fused to a sequence which encodes a functional activation domain in the microorganism. The fused sequences are then introduced into the  
15 microorganism. It is possible that the sequence from the cDNA library may already encode a functional activation domain, for example as described herein in Example 1.

Microorganisms which express the reporter gene are then selected. Since only those microorganisms which express a fusion protein which  
20 includes a binding domain for the target DNA regulatory sequence and an activation domain will stimulate expression of the reporter gene, expression of the reporter gene indicates expression of such a fusion protein.

The gene sequence from the cDNA library introduced into the microorganism which stimulates expression of the reporter gene is then  
25 identified.

According to the above method, the target DNA regulatory sequence preferably includes the subsequence CCG and more preferably includes the subsequence CCGAC.

The "one-hybrid" strategy described in Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993) and used in Example 1 to screen *Arabidopsis* cDNA is an example of this method. This method can be used to screen any plant species for cDNAs that encode a target regulatory sequence, such as a C-repeat/DRE regulatory sequence. According to the "one hybrid" strategy,  
30 yeast strains are constructed that contain a *lacZ* reporter gene with either wild-



type or mutant versions of target regulatory sequences in place of the normal UAS (upstream activator sequence) of the *GAL1* promoter. Yeast strains carrying these reporter constructs produce low levels of  $\beta$ galactosidase and form white colonies on filters containing X-gal. Reporter strains carrying wild-type target regulatory sequences are transformed with a cDNA expression library that contains random cDNA inserts fused to the acidic activator domain of the yeast GAL4 transcription factor "GAL4-ACT". Recombinant plasmids in the expression library that contain a cDNA insert encoding a C-repeat/DRE binding domain fused to GAL4-ACT will express fusion proteins which bind upstream of the *lacZ* reporter genes carrying the wild-type target regulatory sequence, activate transcription of the *lacZ* gene, and result in yeast forming blue colonies on X-gal-treated filters. Alternatively, the sequence from the cDNA library introduced into the microorganism may, as was observed in Example 1, include a sequence encoding an activator domain and thus not utilize the acidic activator domain of the yeast GAL4 transcription factor "GAL4-ACT".

Recombinant plasmids from such "blue yeast" are then isolated and transformed back into reporter strains that contain either a wild-type or mutant version of target regulatory sequence fused to the *lacZ* gene. The plasmids that are desired are those that turn the former strains blue, but not the later, indicating that the cloned DNA binding domain is specific for the target regulatory sequence.

Based on presence of an AP2 binding domain in CBF1, CBF2 and CBF3, Applicants believe that an AP2 binding domain is present in a significant number of the environmental stress tolerance regulatory binding proteins. Accordingly, it is believed that the specificity of the above method for screening for gene sequences encoding a regulatory binding protein can optionally be improved by first selecting cDNA from a plant genome library which includes a potential AP2 domain site. This can be routinely done by selecting probes for selecting sequences in the library which include potential AP2 domain sequences.

#### 4. Screening For Expression Of Environmental Stress Tolerance Regulatory Protein

Once one or more microorganisms are selected which are believed to express a protein capable of binding to the target regulatory element and activate expression of the reporter gene, further analysis can be performed to identify and isolate full length cDNAs; i.e. cDNAs that encode the entire protein that binds to the target regulatory sequence. The coding sequence for the protein can then be cloned into an expression vector, such as the pET bacterial expression vectors (Novagen), and used to produce the protein at high levels. The protein can then be analyzed by gel retardation experiments (See Example 1F) to confirm that it binds specifically to the target regulatory sequence.

Potential sequences can be further screened using known regulatory gene sequences, such as CBF1, CB2, and CBF3, or the presence of an AP2 domain which is believed to be common to a significant class of these genes. Once identified, particular sequences can be transformed into yeast to test for activation of expression of a reporter gene, for example as described in Example 1E.

#### 5. Screening For Binding To Target Regulatory Sequence

Once a regulatory gene sequence is identified, the sequence can be introduced into a microorganism in order to express the protein encoded by the sequence. A gel shift assay, such as the one described in Example 1F, can then be used to test for *in vitro* binding of the expressed protein to the target DNA regulatory sequence.

Mutagenesis of the target DNA regulatory sequence can also be performed in order to evaluate the binding selectivity of the expressed protein. It is preferred that the expressed protein selectively bind to the target DNA regulatory sequence over related sequences with one or more base differences from the target DNA regulatory sequence. For example, Figure 5 is a photograph of a gel from a shift assay in which CBF1 was shown to selectively bind to the wild-type C-repeat/DRE CCGAC.

## 6. Altering The Environmental Stress Tolerance of a Plant.

The present invention also provides a method for recombinant engineered plants with a new or altered response to one or more environmental stresses.

According to one embodiment, a copy of a gene native to a plant which encodes a binding protein according to the present invention is recombinantly introduced into the plant such that the plant expresses a recombinant binding protein encoded by the recombinant copy of the gene.

According to another embodiment, a non-native gene which encodes a binding protein according to the present invention is recombinantly introduced into a plant such that the plant expresses a recombinant binding protein encoded by the recombinant non-native gene.

According to yet another embodiment, a native or non-native DNA regulatory sequence is recombinantly introduced into a plant such that the recombinant DNA regulatory sequence regulates the expression of one or more environmental stress tolerance genes in the plant. The plant includes a gene which encodes a binding protein capable of binding to the recombinant DNA regulatory sequence.

In yet another embodiment, a native or non-native promoter is recombinantly introduced into a plant such that the recombinant promoter regulates the expression of a binding protein which binds to a DNA regulatory sequence.

According to each of the above embodiments, unless otherwise specified, the gene encoding the binding protein, the promoter promoting the expression of the binding protein, the DNA regulatory sequence, and the environmental stress tolerance genes may be non-recombinant or recombinant sequences. The recombinant sequences may be native to the plant or may be non-native to the plant. All the above permutations are intended to fall within the scope of the present invention.

As an example, many plants increase in freezing tolerance in response to low non-freezing temperatures, a process known as cold acclimation. A large number of biochemical changes occur during cold acclimation including the activation of *COR* (Cold Regulated) genes. These genes, which are also

expressed in response to dehydration (e.g., drought and high salinity), are thought to help protect plant cells against the potentially deleterious effects of dehydration associated with freezing, drought and high salinity stress. Indeed, expression of the *COR15a* gene in plants grown at normal temperatures (22 °C) enhances the freezing tolerance of chloroplasts.

By manipulating the expression of *COR* genes, the stress tolerance of crop and horticultural plants could be improved, e.g., engineer broader climate ranges; target stress resistance to stress-sensitive parts of plants; render plants stress-resistant when a stress condition (frost and drought) is imminent. To bring about these effects, however, the expression of the *COR* genes must be manipulated. The gene, *CBF1*, that encodes the transcription factor that binds to the C-repeat/DRE regulatory element present in the promoters of all *COR* genes described to date has been isolated. *CBF1* in yeast activates expression of reporter genes that have been fused to the C-repeat/DRE element. Further, expression of *CBF1* in plants has been shown to activate the expression of *COR* genes.

By introducing modified versions of sequences encoding regulatory binding proteins, such as *CBF1*, into plants, the expression of *COR* genes can be modified, and thereby enhance the freezing and dehydration tolerance of plants.

In each of the above embodiments, expression of the recombinant copy of the regulatory gene may be under the control of a promoter. The promoter may be recombinant or non-recombinant. In the case of recombinant promoters, the promoter may be native or non-native to the plant.

When a recombinant promoter is used, the promoter can be selected to cause expression of the binding protein in a manner which is different than how the binding protein is expressed by the plant in its native state. For example, the promoter may increase the level at which the binding protein is expressed, express the binding protein without being induced by an environmental stress and/or express the binding protein in response to a different form or degree of environmental stress than would otherwise be needed to induce expression of the binding protein. The promoter may also be inducible by an exogenous agent. For example, a strong constitutive promoter could be used to cause increased levels of *COR* gene expression in both non-stress and stressed

plants which in turn, results in enhanced freezing and dehydration tolerance. Examples of such strong constitutive promoters include but are not limited to the nopaline synthase (NOS) and octopine synthase (OCS) promoters, the cauliflower mosaic virus (CaMV) 19S and 35S (Odell et al., Nature **313**: 810-812 (1985)) promoters or the enhanced CaMV 35S promoters (Kay et al., Science **236**: 1299-1302 (1987)).

A tissue-specific promoter could also be used to alter COR gene expression in tissues that are highly sensitive to stress such as embryos in the seed, flower and fruit, thereby enhancing the stress tolerance of these tissues.

Embryo -active promoters include promoters such as the B. napus napin promoter (US Patent 5,420,034), the soybean 7S promoter, the Arabidopsis 12S globulin (cruiferin) promoter (Pang, et al. Plant Molecular Biology **11**:805-820 (1988)), or the maize globulin promoter (Kriz et al. Plant Physiol. **91**:636 (1989); US Patent No. 5,773,691) for use in cereal embryos.

Promoters useful in expressing foreign genes in fruits (Cordes et al., Plant Cell **1**:1025-1034 (1989); Deikman and Fischer, EMBO J. **7**: 3315-3320 (1988); Della Penna et al., Proc. Natl. Acad. Sci. USA **83**: 6420-6424 (1986)) could also be used to alter COR gene expression in fruits. Examples include, but are not limited to, the fruit-specific promoter that was used to express an ADP glucose pyrophosphorylase gene in order to increase the solid content of tomato fruit (Kishore, PCT App. WO 91/19806), the promoter from the 2A11 genomic clone (Pear, et al. Plant Mol. Biol. **13**: 639-651 (1989); US Patent No. 4,943,674) that can be used to control expression of ADP glucose pyrophosphorylase in tomato fruit, the E4 and E8 promoters (Deikman, et al., EMBO J. **7**: 3315-3320 (1988); US Patent No. 5,545,815), the promoter for polygalacturonase, the raspberry fruit promoter described in US Patent No. 5,783,393, fruit-active promoters such as the E8 promoter from tomatoes, and citrus fruit-specific or fruit-active promoters that can be isolated from the CitMT45 cDNA (Moriguchi et al., Gene **12**: 221-227 (1998)) and pSPS2 (Komatsu et al., Mol. Gen. Genet. **252**:346-351 (1996)).

Promoters known to be expressed in developing flowers, particularly in the carpel or pistil tissues, could also be used to alter COR gene expression in flowers. Examples of such promoters include the DefH9 promoter that was used to make parthenocarpic plants and is expressed in the petals, stamens,

carpels and developing ovules (Rotino et al. Nat Biotechnol **15**:1398-401(1997)), the SK2-promoter that was shown to express in the pistil (Ficker et al., Plant Mol Biol **35**: 425-31 (1997)), and the Agamous promoter and intergenic region that was used to express in early and late flowers, and in the inner two whorls of flowers (Sieburth and Meyerowitz, Plant Cell **9**: 355-65 (1997)).

Other tissue-specific promoters that could be used to alter COR gene expression in specific tissues include, but are not limited to, seed-specific promoters for the B. napus napin gene (U.S. Patent No. 5,420,034), the soybean 7S promoter, the Arabidopsis 12S globulin (cruiferin) promoter (Pang, et al. Plant Molecular Biology **11**: 805-820 (1988)), the maize 27kd zein promoter, the rice glutelin 1 promoter and the phytohemagglutinin gene, tuber-specific promoters such as the patatin promoter, and the promoter for the small subunit of ribulose-1,5-bis-phosphate carboxylase (ssRUBISCO) whose expression is activated in photosynthetic tissues such as leaves. It should be noted that other promoters that are known or found to cause specific expression in flowers, seeds or fruits of plants or express in these or other tissues of the plants to cause transcription in plant cells could also be used to alter COR gene expression in the specific tissues according to the present invention.

Altering COR gene expression in specific tissues of plants such as flowers, fruits or seeds may increase frost tolerance of these tissue and prolong the growing seasons for plants. Examples of the specific tissues of plants according to the present invention include, but not limited to, frost-resistant flowers in strawberries, peaches, blueberries, cherries, apricots, daffodils, apples, and plums; frost-resistant canola or rape seeds for preventing the formation of green seeds at harvest; frost-resistant barley seeds for maintaining malting ability; and frost-resistant fruits including true berries such as tomato, grape, blueberry, cranberry, currant, and eggplant; stone fruits (drupes) such as cherry, plum, apricot, peach, nectarine and avocado; and compound fruits (drupelets) such as raspberry and blackberry; in citrus fruits such as oranges, lemons, grapefruit and tangerines; and in melons such watermelon, cantaloupe, honeydew, cucumber, and squash.

In addition, the COR gene expression can also be altered in specific tissues of the following plants according to the present invention: cauliflower (Brassica oleracea), artichoke (Cynara scolymus), fruits such as apple (Malus, e.g. domesticus), banana (Musa, e.g. acuminata), berries (such as the currant, Ribes, e.g. rubrum), cherries (such as the sweet cherry, Prunus, e.g. avium), cucumber (Cucumis, e.g. sativus), grape (Vitis, e.g. vinifera), lemon (Citrus limon), melon (Cucumis melo), nuts (such as the walnut, Juglans, e.g. regia; peanut, Arachis hypogaeae), orange (Citrus, e.g. maxima), peach (Prunus, e.g. persica), pear (Pyrus, e.g. communis), pepper (Solanum, e.g. capsicum), plum (Prunus, e.g. domestica), strawberry (Fragaria, e.g. moschata), tomato (Lycopersicon, e.g. esculentum), leafs, such as alfalfa (Medicago, e.g. sativa), cabbages (such as Brassica oleracea), endive (Cichoreum, e.g. endivia), leek (Allium, e.g. porrum), lettuce (Lactuca, e.g. sativa), spinach (Spinacia e.g. oleraceae), tobacco (Nicotiana, e.g. tabacum), roots, such as arrowroot (Maranta, e.g. arundinacea), beet (Beta, e.g. vulgaris), carrot (Daucus, e.g. carota), cassava (Manihot, e.g. esculenta), turnip (Brassica, e.g. rapa), radish (Raphanus, e.g. sativus), yam (Dioscorea, e.g. esculenta), sweet potato (Ipomoea batatas) and seeds, such as bean (Phaseolus, e.g. vulgaris), pea (Pisum, e.g. sativum), soybean (Glycin, e.g. max), wheat (Triticum, e.g. aestivum), barley (Hordeum, e.g. vulgare), corn (Zea, e.g. mays), rice (Oryza, e.g. sativa), tubers such as kohlrabi (Brassica, e.g. oleraceae), and potato (Solanum, e.g. tuberosum).

Alternatively, an inducible promoter may be used to control the expression of the regulatory binding protein, such as CBF1, in plants. Because, in some cases, constitutive expression of higher levels of CBF proteins may have some detrimental effects on plant growth and development, the controlled expression of CBF genes is especially advantageous. For example, a promoter could be used to induce the expression of CBF proteins only at a proper time, such as prior to a frost that may occur earlier or later in the growing season of a plant, thereby prolonging the growing season of a crop and increasing the productivity of the land. This may be accomplished by applying an exogenous inducer by a grower whenever desired. Alternatively, a promoter could be used which turns on at a temperature that is warmer than

the temperature at which the plant normally exhibits cold tolerance. This would enable the cold tolerance thermostat of a plant to be altered. Similarly, a promoter can be used which turns on at a dehydration condition that is wetter than the dehydration condition at which the plant normally exhibits dehydration tolerance. This would enable the level at which a plant responds to dehydration to be altered.

Promoters which are known or are found to cause inducible transcription of the DNA into mRNA in plant cells can be used in the present invention. Such promoters may be obtained from a variety of sources such as plant and inducible microbial sources, and may be activated by a variety of exogenous stimuli, such as cold, heat, dehydration, pathogenesis and chemical treatment. The particular promoter selected is preferably capable of causing sufficient expression of the regulatory binding protein, such as CBF1, to enhance plant tolerance to environmental stresses. Examples of promoters which may be used include, but are not limited to, the promoter for the DRE (C-repeat) binding protein gene *dreb2a* (Liu, et al. Plant Cell **10**: 1391-1406 (1998)) that is activated by dehydration and high-salt stress, the promoter for delta 1-pyrroline-5-carboxylate synthetase (P5CS) whose expression is induced by dehydration, high salt and treatment with plant hormone abscisic acid (ABA) (Yoshida, et al., Plant J. **7** 751-760 (1987)), the promoters for the *rd22* gene from *Arabidopsis* whose transcription is induced under by salt stress, water deficit and endogenous ABA (Yamaguchi-Shinozaki and Shinozaki, Mol Gen Genet **238** 17-25 (1993)), the promoter for the *rd29b* gene (Yamaguchi-Shinozaki and Shinozaki, Plant Physiol., **101** 1119-1120 (1993)) whose expression is induced by desiccation, salt stress and exogenous ABA treatment (Ishitani et al., Plant Cell **10** 1151-1161 (1998)), the promoter for the *rab18* gene from *Arabidopsis* whose transcripts accumulate in plants exposed to water deficit or exogenous ABA treatment, and the promoter for the pathogenesis-related protein 1a (PR-1a) gene whose expression is induced by pathogenesis organisms or by chemicals such as salicylic acid and polyacrylic acid.

It should be noted that the promoters described above may be further modified to alter their expression characteristics. For example, the drought/ABA inducible promoter for the *rab18* gene may be incorporated into



seed-specific promoters such that the rab18 promoter is drought/ABA inducible only when developing seeds. Similarly, any number of chimeric promoters can be created by ligating a DNA fragment sufficient to confer environmental stress inducibility from the promoters described above to constitute promoters with  
5 other specificities such as tissue-specific promoters, developmentally regulated promoters, light-regulated promoters, hormone-responsive promoters, etc. This should result in the creation of chimeric promoters capable of being used to cause expression of the regulatory binding proteins in any plant tissue or combination of plant tissues. Expression can also be made to occur either at a  
10 specific time during a plant's life cycle or throughout the plant's life cycle.

According to the present invention, an expression vector can be constructed to express the regulatory binding protein in the transformed plants to enhance their tolerance to environmental stresses. In one embodiment, the DNA construct may contain (1) an inducible promoter that activates expression  
15 of the regulatory binding protein in response to environmental stimuli; (2) a sequence encoding the regulatory binding protein; and (3) a 3' non-translated region which enables 3' transcriptional termination and polyadenylation of the mRNA transcript. The inducible promoter may be any one of the natural or recombinant promoters described above. The gene encoding the regulatory  
20 binding protein can be any one disclosed in the present invention. The 3' region downstream from this gene should be capable of providing a polyadenylation signal and other regulatory sequences that may be required for the proper expression and processing of a mRNA may be operably linked to the 3' end of a structural gene to accomplish the invention. This may include the native 3' end  
25 of the homologous gene from which the regulatory binding protein and/or the inducible promoter is derived, the 3' end from a heterologous gene encoding the same protein from other species, the 3' end from viral genes such as the 3' end of the 35S or the 19S cauliflower mosaic virus transcripts, the 3' end of the opine synthesis genes of *Agrobacterium tumefaciens*, or the 3' end sequences  
30 from any source such that the sequence employed provides the necessary regulatory information within its nucleic acid sequence to result in the proper expression of the promoter/coding region combination to which the 3' end sequence is operably linked.

A variety of expression vectors can be used to transfer the gene encoding the regulatory binding protein as well as the desired promoter into the plant. Examples include but not limited to those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as well as those disclosed by Herrera-Estrella, L., et al., *Nature* **303**: 209(1983), Bevan, M., *Nucl. Acids Res.* **12**: 8711-8721 (1984), Klee, H. J., *Bio/Technology* **3**: 637-642 (1985), and EPO Publication 120,516 (Schilperoort et al.) for dicotyledonous plants. Alternatively, non-Ti vectors can be used to transfer the DNA constructs of this invention into monotyledonous plants and plant cells by using free DNA delivery techniques. Such methods may involve, for example, the use of liposomes, electroporation, microprojectile bombardment, silicon carbide whiskers, viruses and pollen. By using these methods transgenic plants such as wheat, rice (Christou, P., *Bio/Technology* **9**: 957-962 (1991)) and corn (Gordon-Kamm, W., *Plant Cell* **2**: 603-618 (1990)) are produced. An immature embryo can also be a good target tissue for monocots for direct DNA delivery techniques by using the particle gun (Weeks, T. et al., *Plant Physiol.* **102**: 1077-1084 (1993); Vasil, V., *Bio/Technology* **10**: 667-674 (1993); Wan, Y. and Lemeaux, P., *Plant Physiol.* **104**: 37-48 (1994), and for *Agrobacterium*-mediated DNA transfer (Hiei et al., *Plant J.* **6**: 271-282 (1994); Rashid et al., *Plant Cell Rep.* **15**: 727-730 (1996); Dong, J., et al., *Mol. Breeding* **2**: 267-276 (1996); Aldemita, R. and Hodges, T., *Planta* **199**: 612-617 (1996); Ishida et al., *Nature Biotech.* **14**: 745-750 (1996)).

In one embodiment, the plasmid vector pMEN020 is preferred, which is derived from a Ti plasmid pMON10098 which is the type of binary vector described in U.S. Patent Nos. 5,773,701 and 5,773,696. PMEN20 differs from pMON10098 by the substitution of a KpnI, Sall, SacI, SacII, NotI, and XbaI restriction sites between the ECaMV 35S promoter and the E9 3' region. Plasmid pMON10098 contains the following DNA segments. Starting at the bottom of the plasmid map is the origin of bacterial replication for maintenance in *E. coli* (ori-322). Moving in a counter-clockwise direction on the map, next is ori-V, which is the vegetative origin of replication (Stalker et al. *Mol. Gen. Genet.* **181**:8-12 (1981)). Next is the left border of the T-DNA. Next is the chimeric gene used as the selectable marker. The chimera includes the 0.35 kilobase (kb) of the cauliflower mosaic virus 35S promoter (P-35S) (Odell et al. (1985) *Nature* **313**:810-812). , a 0.84 kb neomycin phosphotransferase type II

gene (KAN) and a 0.25 kb 3' non-translated region of the nopaline synthase gene (NOS 3') (Fraley et al. (1983) Proc. Natl. Acad. Sci. USA 80:1803-1807). The next sequence contains the enhanced CaMV 35S promoter and E9 3' region gene cassette and restriction sites for inserting genes such as the coding region of CBF genes. This chimeric gene cassette ends with the 0.65 kb of the E9 3' region from the pea small subunit of RUBISCO gene (U.S. Patent No. 5,773,701). Next is the right border of the T-DNA. Next is the 0.93 kb fragment isolated from transposon Tn7 that encodes the bacterial spectinomycin/streptomycin resistance (Spc/Str), which is a determinant for selection in *E. coli* and *Agrobacterium tumefaciens* (Fling et al., *Nucl. Acids Res.* 13:7095-7106 (1985)).

The pMEN020 plasmid construct is a binary cloning vector that contains both *E. coli* and *Agrobacterium tumefaciens* origins of DNA replication but no *vir* genes encoding proteins essential for the transfer and integration of the target gene inserted in the T-DNA region. PMEN020 requires the *trfA* gene product to replicate in *Agrobacterium*. The strain of *Agrobacterium* containing this *trfA* gene is called the ABI strain and is described below and in U.S. Patent Nos. 5,773,701 and 5,773,696. This cloning vector serves as an *E. coli*-*Agrobacterium tumefaciens* shuttle vector. All of the cloning steps are carried out in *E. coli* before the vector is introduced into ABI strain of *Agrobacterium tumefaciens*.

In another embodiment, pMEN050 is preferred, which is derived from pMEN020 by replacing the NptII kanamycin resistance gene with the Bar gene (US Patent No. 5,646,024) by using the same cloning method described above for pMEN020.

The recipient ABI strain of *Agrobacterium* carries a modified defective Ti plasmid that serves as a helper plasmid containing a complete set of *vir* genes but lacks portions or all of the T-DNA region. ABI is the A208 *Agrobacterium tumefaciens* strain carrying the disarmed pTiC58 plasmid pMP90RK (Koncz et al. *Mol. Gen. Genet.* 204:383-396 (1986)). The disarmed Ti plasmid provides the *trfA* gene functions that are required for autonomous replication of the binary vectors after transfer into the ABI strain. When plant tissue is incubated with the ABI::binary vector strains, the vectors are transferred to the plant cells by the *vir* functions encoded by the disarmed pMP90RK Ti plasmid. After the

introduction of the binary vector into the recipient *Agrobacterium*, the *vir* gene products mobilize the T-DNA region of the pMEN020 plasmid to insert the target gene, e.g. the gene encoding the regulatory binding protein, into the plant chromosomal DNA, thus transforming the cell.

5           It should be noted that methods for transforming a wide variety of different dicots and obtaining transgenic plants are well documented in the literature (See Gasser and Fraley *Science* **244**:1293 (1989); Fisk and Dandekar, *Scientia Horticulturae* **55**: 5-36 (1993); Christou *Agro Food Industry Hi Tech* March/April: p.17 (1994), and the references cited therein).

10           Methods for producing transgenic plants among the monocots are also available. Successful transformation and plant regeneration have been achieved in asparagus (*Asparagus officinalis*; Bytebier et al. *Proc. Natl. Acad. Sci. USA* **84**:5345 (1987)); barley (*Hordeum vulgare*; Wan and Lemaux, *Plant Physiol* **104**:37 (1994)); maize (*Zea mays*; Gordon-Kamm et al., *Plant Cell* **2**:603 (1990); Fromm et al. *Bio/Technology* **8**:833 (1990); Koziel et al. *Bio/Technology* **11**: 194 (1993)); oats (*Avena sativa*, Somers et al. *Bio/Technology* **10**: 1589 (1992)); orchardgrass (*Dactylis glomerata*; Horn et al. *Plant Cell Rep.* **7**: 469 (1988)); rice (*Oryza sativa*, including indica and japonica varieties; Toriyama et al. *Bio/Technology* **6**:10 (1988); Zhang et al. *Plant Cell Rep.* **7**: 379 (1988); Luo and Wu *Plant Mol. Biol. Rep.* **6**:165 (1988); Zhang and Wu, *Theor. Appl. Genet.* **76**: 835 (1988); Christou et al. *Bio/Technology* **9**: 957 (1991); rye (*Secale cereale*; De la Pena et al. *Nature* **325**: 274 (1987)); sorghum (*Sorghum bicolor*; Cassas et al. *Proc. Natl. Acad. Sci. USA* **90**:11212 (1993)); sugar cane (*Saccharum spp.*; Bower and Birch *Plant J.* **2**: 409 (1992)); tall fescue (*Festuca arundinacea*; Wang et al. *Bio/Technology* **10**:691 (1992)); turfgrass (*Agrostis palustris*; Zhong et al. *Plant Cell Rep.* **13**:1 (1993)); wheat (*Triticum aestivum*; Vasil et al. *Bio/Technology* **10**: 667 (1992); Troy Weeks et al. *Plant Physiol.* **102**:1077 (1993); Becker et al. *Plant J.* **5**:299 (1994)).

30           After transformation of cells or protoplasts, the choice of methods for regenerating fertile plants is not particularly important. Suitable protocols are available for *Leguminosae* (alfalfa, soybean, clover, etc.), *Umbelliferae* (Carrot, celery, parsnip), *Cruciferae* (cabbage, radish, rapeseed, broccoli, etc.), *Curcubitaceae* (melons and cucumber), *Gramineae* (wheat, corn, rice, barley,

millet, etc.), *Solanaceae* (potato, tomato, tobacco, peppers, etc.), and various other crops See protocols described in Ammirato et al. (1984) *Handbook of Plant Cell Culture –Crop Species*. Macmillan Publ. Co. Shimamoto et al. *Nature* **338**:274-276 (1989); Fromm et al., *Bio/Technology* **8**:833-839 (1990);  
5 Vasil et al. *Bio/Technology* **8**:429-434 (1990).

It is envisioned that the present invention can be used to introduce, change and/or augment the environmental stress tolerance of a plant by introducing and causing the expression of environmental stress tolerance in a manner which the plant does not exhibit in its native form. For example, by  
10 using different promoters in combination with recombinant regulatory genes, native environmental stress tolerance genes can be expressed independent of environmental stress, made responsive to different levels or types of environmental stress, or rendered inducible independent of an environmental stress. Further, selection of the promoter can also be used to determine what  
15 tissues in the plant express the binding protein as well as when the expression occurs in the plant's lifecycle. By selecting a promoter which regulates in what tissues and when in a plant's life the promoter functions to regulate expression of the binding protein, in combination with the selecting how that promoter regulates expression (level of expression and/or type of environmental or  
20 chemical induction), an incredible range of control over the environmental stress responses of a plant can be achieved using the present invention.

By recombinantly introducing a native environmental stress tolerance gene into a plant in combination with a recombinant regulatory gene under the control of an inducible promoter, a plant can be engineered which includes its  
25 native environmental stress tolerance as well as inducible environmental stress tolerance. This might be useful for inducing a cold stress tolerance reaction in anticipation of a frost.

By recombinantly introducing a non-native environmental stress tolerance gene into a plant in combination with a recombinant regulatory gene,  
30 a plant can be engineered which includes environmental stress tolerance properties that the plant would not otherwise have. In this regard, plants from warmer climates can be engineered to include one or more cold tolerance genes along with a regulatory gene needed to cause expression of the cold tolerance genes in the plant so that the engineered plant can survive better in a

colder climate. Similarly, a plant can be engineered to include one or more dehydration tolerance genes along with a regulatory gene needed to cause expression of the dehydration tolerance gene so that the engineered plant can grow better in a dryer climate. In this regard, it should be possible to take a  
5 plant which grows well in a first climate and engineer it to include stress tolerance genes and regulatory genes native to a second climate so that the plant can grow well in the second climate.

By modifying the promoter controlling the expression of the gene encoding a binding protein which regulates the expression of environmental  
10 stress tolerance genes, the operation of native, non-recombinant environmental stress tolerance genes and regulatory genes can be changed. For example, the conditions under which the stress tolerance genes are expressed can be changed. Expression can also be rendered inducible by an exogenous agent.

15 **7. Methods For Detecting Stress Tolerance  
Regulatory Gene Homologs.**

Once one DNA sequence encoding an environmental stress tolerance regulatory binding protein has been identified, several methods are available  
20 for using that sequence and knowledge about the protein it encodes to identify homologs of that sequence from the same plant or different plant species. For example, let us assume that a cDNA encoding a first target binding domain has been isolated from plant species "A." The DNA sequence encoding the first target DNA regulatory sequence could be radiolabeled and used to screen  
25 cDNA libraries of plant species "A," or any other plant species, for DNA inserts that encode proteins related to the first target DNA regulatory sequence. This could be done by screening colony or phage "lifts" using either high ( $T_m$  of about  $-10^{\circ}\text{C}$ ) or low ( $T_m$  of about  $-30^{\circ}\text{C}$  or lower) stringency DNA hybridization conditions (Sambrook, J. et al, Molecular Cloning. A Laboratory  
30 Manual Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 2nd Ed. (1989)). cDNA inserts that hybridize with the first target DNA regulatory sequence could be sequenced and compared to the original first target DNA regulatory sequence. If the insert is confirmed to encode a polypeptide similar to the first target DNA regulatory sequence, the insert could be cloned into an  
35 expression vector to produce the encoded protein. The protein would then be

analyzed by gel retardation experiments to confirm that it binds specifically to the first target DNA regulatory sequence.

It is recognized that not all proteins that bind to a first target DNA regulatory sequence will be transcriptional activators. However, a number of routine tests may be performed in order to determine whether a particular protein is in fact a transcriptional activator. One test involves expressing the protein in yeast strains which contain the target DNA regulatory sequence fused to the *lacZ* reporter gene, as described above. If the protein is a transcriptional activator, it should activate expression of the reporter gene and result in blue colonies.

Another test is a plant transient assay. In this case, a reporter gene, such as GUS, carrying the target DNA regulatory sequence as an upstream activator is introduced into plant cells (e.g. by particle bombardment) with or without a the putative transcriptional activator under control of a constitutive promoter. If the protein is an activator, it will stimulate expression of the reporter (this may be further enhanced if the plant material is placed at low temperature or is subjected to water stress as the C-repeat/DRE is responsive to low temperature and dehydration).

Significantly, once a target DNA regulatory sequence is identified, the sequence can be fused to any potential activator or repressor sequence to modify expression of plant genes that carry the target regulatory sequence as a control element. That is, the DNA regulatory sequence can be used to target "managed" expression of the battery of environmental stress tolerance related genes in a given plant species.

It is possible that the target DNA regulatory sequence of the regulatory element that imparts environmental stress tolerance related gene expression in plant species "A" might be slightly different from the analogous target DNA regulatory element that imparts environmental stress tolerance in species "B." Thus, optimal regulation of the battery of environmental stress tolerance related genes in a given species may require the use of the regulatory binding proteins from that or a closely related plant species. Knowledge of gene sequences which encode for proteins which bind to the DNA regulatory sequence of the regulatory element, in combination with knowledge of the DNA regulatory

sequence, greatly simplify the identification of sequences encoding binding proteins native to the plant species.

With the advent of fast and efficient DNA sequencing technologies, the number of plant genomes recorded on computer databases is growing rapidly. These computer databases can be used to search for homologs to CBF sequences identified in this application as well as other sequences which encode binding proteins which regulate cold tolerance genes. As more and more binding protein sequences are identified and the number of computerized plant genome databases increase, searching computer databases for additional sequences encoding binding proteins which regulate cold tolerance genes will become increasingly simplified.

#### 8. Preparation Of Binding Proteins Derivatives Using Sequences Identified In This Application.

According to the present invention, the binding protein is a protein which is capable of binding to a DNA regulatory sequence which regulates expression of one or more environmental stress tolerance genes in a plant. These DNA regulatory sequences are preferably a member of the CCG family of regulatory sequences and more preferably a member of the CCGAC family of regulatory sequences.

Numerous amino acid sequences for CBF binding protein homologs are disclosed in this application including those shown in Figures 2B, 14, and 18B and listed in SEQ. I.D. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95. Nucleic acid sequences encoding these CBF binding protein homologs are disclosed in this application in Figures 2B, 12, 13, and 18A and listed in SEQ. I.D. Nos. 1, 12, 14, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, and 94. These sequences were derived from a variety of different plant species including *Arabidopsis*, *Brassica juncea*, *Brassica napus*, *Brassica oleracea*, *Brassica rapa*, *Glycine max*, *Raphanus sativus* and *Zea Maize*.

The sequences identified in these figures may generally be divided into three regions: AP2 domain, amino terminus domain, and carboxy terminus domain. Figured 19A-19E show different AP2 domains from these homologs



and consensus sequences between the different AP2 domains shown.

Figure 19A shows an amino acid alignment of the AP2 domains of several CBF proteins with the consensus sequence between the proteins highlighted as well as a comparison of the AP2 domains with that of the tobacco DNA binding protein EREBp2. Figure 19B shows an amino acid alignment of the AP2 domains of several CBF proteins including dreb2a and dreb2b with the consensus sequence between the proteins highlighted. Figure 19C shows an amino acid alignment of the AP2 domains of several CBF proteins including dreb2a, dreb2b, and tiny with the consensus sequence between the proteins highlighted. Figure 19D shows a consensus sequence corresponding to the difference between the consensus sequence shown in Figures 19A and tiny. Figure 19E shows a consensus sequence corresponding to the difference between the consensus sequence shown in Figures 19B and tiny.

Figures 21A and 21B show different carboxy terminus domains from these homologs and consensus sequences between the different carboxy terminus domains shown.

The binding proteins utilized in the present invention include classes of binding proteins which satisfy one or more of the following requirements:

the binding protein comprises an AP2 domain which comprises a consensus sequence sufficiently homologous to any one of the consensus sequences shown in Figures 19A, 19B, or 19C that the binding protein is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises an AP2 domain which comprises a consensus sequence shown in Figures 19A, 19B or 19C;

the binding protein comprises an AP2 domain which comprises the amino acid residues shown in Figures 19D or 19E;

the binding protein comprises an AP2 domain which is sufficiently homologous to at least one of the AP2 domains shown in the application such that it is capable of binding to a CCG regulatory sequence, preferably a CCGAC regulatory sequence;

the binding protein comprises one of the AP2 domain sequences shown in this application, including, but not limited to SEQ. I.D. Nos. 2, 13, 15, 39, 41,

43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95;

5 the binding protein comprises a sequence which comprises one of the amino terminus domains shown in Figure 20 (it is noted that the sequence need not be at the amino terminus of the binding protein);

the binding protein comprises the consensus sequence for the amino terminus domains shown in Figure 20, (it is noted that the sequence need not be at the amino terminus of the binding protein);

10 the binding protein comprises a sequence which comprises one of the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein);

the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21A (it is noted that the sequence need not be at the carboxy terminus of the binding protein); and

15 the binding protein comprises the consensus sequence for the carboxy terminus domains shown in Figure 21B (it is noted that the sequence need not be at the carboxy terminus of the binding protein).

20 The sequence of the binding protein may be a naturally occurring sequence such as the ones shown in SEQ. ID. Nos. 2, 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95 or may be a non-naturally occurring sequence. It is noted, however, that binding proteins according to the present invention are intended to encompass non-naturally occurring sequences which are derivatives of the classes of binding proteins taught herein.

25 Additional binding proteins may be constructed using one of the AP2 domains taught herein or the consensus sequence of these AP2 domains. It may be desirable to include with the AP2 domain a transcription activation region. The transcription activation region may be native to the plant or non-native to the plant in which the binding protein will be used. For example, the  
30 sequence may include a subsequence which encodes a binding domain for the DNA regulatory sequence fused to a transcription activating region, such as the transcription activating region of VP16 or GAL4. Optionally, one can include in the binding protein one of the amino terminus domains, the consensus sequence for the amino terminus domain, one of the carboxy terminus domains

and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

5            Optionally, the binding protein can be viewed as comprising one of the amino terminus domains, the consensus sequence for the amino terminus domain, one of the carboxy terminus domains and/or the consensus sequence for the carboxy terminus domains. It is noted that the amino terminus domain may be positioned away from the amino terminus of the new binding protein  
10           and the carboxy terminus domain may be positioned away from the carboxy terminus of the new binding protein.

## EXAMPLES

### 1. Isolation and Analysis of *Arabidopsis Thaliana* cDNA Clone (CBF1) Encoding C-repeat/DRE Binding Factor

The following example describes the isolation of an *Arabidopsis thaliana* cDNA clone that encodes a C-repeat/DRE binding factor, CBF1 (C-repeat/DRE Binding Factor 1). Expression of CBF1 in yeast was found to activate transcription of reporter genes containing the C-repeat/DRE (CCGAC) as an upstream activator sequence. Meanwhile, CBF1 did not activate transcription of mutant versions of the CCGAC binding element, indicating that CBF1 is a transcription factor that binds to the C-repeat/DRE. Binding of CBF1 to the C-repeat/DRE was also demonstrated in gel shift assays using recombinant CBF1 protein expressed in *Escherichia coli*. Analysis of the deduced CBF1 amino acid sequence indicated that the protein has a potential nuclear localization sequence, a possible acidic activation domain and an AP2 domain, a DNA-binding motif of about 60 amino acids that is similar to those present in *Arabidopsis* proteins APETALA2, AINTEGUMENTA and TINY, the tobacco ethylene response element binding proteins, and numerous other plant proteins of unknown function.

#### A. Materials

**Plant material and cold treatment.** *A. thaliana* (L.) Heyn. ecotype RLD plants were grown in pots in controlled environment chambers at 22 °C under constant illumination with cool-white fluorescent lamps ( 100  $\mu\text{mol m}^{-2}\text{s}^{-1}$ ) essentially as described (Gilmour, S. J., Plant Physiol. **87**:745-750 (1988)). Plants were cold-treated by placing pots in a cold room at 2.5 °C under constant illumination with cool-white fluorescent lamps ( 25  $\mu\text{mol m}^{-2}\text{s}^{-1}$ ) for the indicated times.

**Arabidopsis cDNA expression library.** The *Arabidopsis* pACT cDNA expression library was constructed by John Walker and colleagues (NSF/DOE/USDA Collaborative Research in Plant Biology Program grant USDA 92-37105-7675) and deposited in the *Arabidopsis* Biological Resource Center (stock #CD4-10).

**Yeast reporter strains.** Oligonucleotides (Table 1) (synthesized at the MSU Macromolecular Structure Facility) encoding either wild-type or mutant versions of the C-repeat/DRE were ligated into the *Bgl*II site of the *lacZ* reporter vector pBgl-lacZ (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993);  
 5 kindly provided by Joachim Li). The resulting reported constructs were integrated into the *ura3* locus of *Saccharomyces cerevisiae* strain GGY1 (*MAT gal4 gal80 ura3 leu2 his3 ade2 tyr*) (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993); provided by Joachim Li) by transformation and selection for uracil prototrophy.

10 **E. coli strains.** *Escherichia coli* strain GM2163 containing plasmid pEJS251 was deposited under the Budapest Treaty on May 17, 1996 with the American Type Culture Collection, Rockville, Maryland as ATCC 98063. It is available by name and number pursuant to the provisions of the Budapest Treaty.

15

**TABLE 1**

Oligonucleotides encoding wild type and mutant versions of the C-repeat/DRE

Oligonucleotide	C-repeat /DRE*	Sequence	SEQ ID NO:
MT50	<i>COR15a</i>	GatcATTTTCATGGCCGACCTGCTTTTT	3
MT52	M1 <i>COR15a</i>	CACAATTTCAaGaattcaCTGCTTTTTT	4
MT80	M2 <i>COR15a</i>	GatcATTTTCATGGtatgtCTGCTTTTTT	5
MT125	M3 <i>COR15a</i>	GatcATTTTCATGGaatcaCTGCTTTTTT	6
MT68	<i>COR15b</i>	GatcACTTGATGGCCGACCTCTTTTTT	7
MT66	<i>COR78-1</i>	GatcAATATACTACCGACATGAGTTCT	8
MT86	<i>COR78-2</i>	ACTACCGACATGAGTTCCAAAAAGC	9

\*The C-repeat/DRE sequences tested are either wild-type found in the  
 20 promoters of *COR15a* (Baker, S. S., et al., Plant. mol. Biol. **24**:701-713 (1994)), *COR15b* or *COR78/RD29A* (Horvath, D. P., et al., Plant Physiol. **103**:1047-1053 (1993); Yamaguchi-shinozaki, K., et al., The Plant Cell **6**:251-264 (1994)) or are mutant versions of the *COR15a* C-repeat/DRE (M1*COR15a*, M2*COR15a* and M3*COR15a*).

#Uppercase letters designate bases in wild type C-repeat/DRE sequences. The core CCGAC sequence common to the above sequences is indicated in bold type. Lowercase letters at the beginning of a sequence indicate bases added to facilitate cloning. The lowercase letters that are underlined indicate the mutations in the C-repeat/DRE sequence of *COR15a*.

## B. Methods

10                   **Screen of Arabidopsis cDNA library.** The Arabidopsis pACT cDNA expression library was screened for clones encoding C-repeat/DRE environmental stress response regulatory elements by the following method. The cDNA library, harbored in *Escherichia coli* BNN132, was amplified by inoculating 0.5 ml of the provided glycerol stock into 1 L of M9 minimal glucose  
15                   medium (Sambrook, J. et al, Molecular Cloning. A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY., 2nd Ed. (1989)) and shaking the bacteria for 20 h at 37 °C. Plasmid DNA was isolated and purified by cesium chloride density gradient centrifugation (Sambrook et al (1989)) and transformed into the yeast GGY1 reporter strains selecting for leucine  
20                   prototrophy. Yeast transformants that had been grown for 2 or 3 days at 30 °C were overlaid with either a nitrocellulose membrane filter (Schleicher and Schuell, Keene, NH) or Whatman #50 filter paper (Hillsboro, OR) and incubated overnight at 30 °C. The yeast impregnated filters were then lifted from the plate and treated with X-gal (5-bromo-4-chloro-3-indolyl -D-  
25                   galactosidase) to assay colonies for beta-galactosidase activity (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993)). Plasmid DNA from "positive" transformants (those forming blue colonies on the X-gal-treated filters) was recovered (Strathern, J. N., and D. R. Higgins, Methods Enzymol. **194**:319-329 (1991)), propagated in *E. coli* DH5 and transformed back into the yeast  
30                   reporter strains to confirm activity.

**Yeast transformation and quantitative beta-galactosidase assays.** Yeast were transformed by either electroporation (Becker, D. M., et al., Methods Enzymol. **194**:182-187 (1991)) or the lithium acetate/carrier DNA

method (Schiestl, R. H., et al., Current Genetics 16:339-346 (1989)).  
Quantitative *in vitro* beta-galactosidase assays were done as described (Rose, M., et al., Methods Enzymol. 101:167-180 (1983)).

5                   **Expression of CBF1 protein in *E. coli* and yeast.** CBF1 was  
expressed in *E. coli* using the pET-28a(+) vector (Novagen, Madison, WI). The  
*Bgl*II-*Bcl*II restriction fragment of pACT-11 encoding CBF1 was ligated into the  
*Bam*HI site of the vector bringing CBF1 under control of the T7 phage  
promoter. The construct resulted in a "histidine tag," a thrombin recognition  
10                   sequence and a "T7 epitope tag" being fused to the amino terminus of CBF1.  
The construct was transformed into *E. coli* BL21 (DE3) and the recombinant  
CBF1 protein was expressed as recommended by the supplier (Novagen).  
Expression of CBF1 in yeast was accomplished by ligating restriction  
fragments encoding CBF1 (the *Bcl*II-*Bgl*II and *Bgl*II-*Bgl*II fragments from pACT-  
15                   11) into the *Bgl*II site of pDB20.1 (Berger, S. L., et al., Cell 70:251-265 (1992);  
kindly provided by Steve Triezenberg) bringing CBF1 under control of the  
constitutive *ADC1* (alcohol dehydrogenase constitutive 1) promoter.

**Gel shift assays.** The presence of expressed protein which  
20                   binds to a C-repeat/DRE binding domain was evaluated using the following gel  
shift assay. Total soluble *E. coli* protein (40 ng) was incubated at room  
temperature in 10 µl of 1X binding buffer [15 mM HEPES (pH 7.9), 1 mM  
EDTA, 30mM KCl, 5% glycerol, 5% BSA, 1mM DTT) plus 50 ng poly(dI-  
dC):poly(dI-dC) (Pharmacia, Piscataway, NJ) with or without 100 ng competitor  
25                   DNA. After 10 min, probe DNA (1 ng) that was <sup>32</sup>P-labeled by end-filling  
(Sambrook et al, 1989) was added and the mixture incubated for an additional  
10 min. Samples were loaded onto polyacrylamide gels (4% w/v) and  
fractionated by electrophoresis at 150V for 2h (Sambrook et al). Probes and  
competitor DNAs were prepared from oligonucleotide inserts ligated into the  
30                   *Bam*HI site of pUC118 (Vieira, J., et al., Methods Enzymol. 153:3-11 (1987)).  
Orientation and concatenation number of the inserts were determined by  
dideoxy DNA sequence analysis (Sambrook, et al, (1989)). Inserts were  
recovered after restriction digestion with *Eco*RI and *Hind*III and fractionation on  
polyacrylamide gels (12% w/v) (Sambrook et al, 1989).

**Northern and southern analysis.** Northern and southern analysis was performed as follows. Total RNA was isolated from Arabidopsis (Gilmour, S. J., et al., Plant Physiol. **87**:745-750 (1988)) and the poly(A)<sup>+</sup> fraction purified using oligo dT cellulose (Sambrook, et al (1989)). Northern transfers were prepared and hybridized as described (Hajela, R. K., et al., Plant Physiol. **93**:1246-1252 (1990)) except that high stringency wash conditions were at 50 °C in 0.1X SSPE [X SSPE is 3.6 M NaCl, 20 mM EDTA, 0.2 M Na<sub>2</sub>HPO<sub>4</sub> (pH7.7)], 0.5% SDS. Membranes were stripped in 0.1 X SSPE, 0.5% SDS at 95 °C for 15 min prior to re-probing. Total Arabidopsis genomic DNA was isolated (Stockinger, E. J., et al., J. Heredity, **87**:214-218 (1996)) and southern transfers prepared (Sambrook et al 1989) using nylon membranes (MSI, Westborough, MA). High stringency hybridization and wash conditions were as described by Walling et al (Walling, L. L., et al., Nucleic Acids Res. **16**:10477-10492 (1988)). Low stringency hybridization was in 6X SSPE, 0.5% SDS, 0.25% low fat dried milk at 60 °C. Low stringency washes were in 1X SSPE, 0.5% SDS at 50 °C. Probes used for the entire CBF1 coding sequence and 3' end of CBF1 were the *Bcl*I/*Bgl*II and *Eco*RV/*Bgl*II restriction fragments from pACT-11, respectively, that had been gel purified (Sambrook et al (1989)). DNA probes were radiolabeled with <sup>32</sup>P-nucleotides by random priming (Sambrook). Autoradiography was performed using hyperfilm-MP (Amersham, Arlington Heights, IL). Radioactivity was quantified using a Betascope 603 blot analyzer (Betagen Corp., Waltham, MA).

**C. Screen of Arabidopsis cDNA library for sequence encoding a C-repeat/DRE binding domain.**

The "one-hybrid" strategy (Li, J. J. and I. Herskowitz, Science **262**:1870-1874 (1993)) was used to screen for Arabidopsis cDNA clones encoding a C-repeat/DRE binding domain. In brief, yeast strains were constructed that contained a *lacZ* reporter gene with either wild-type or mutant C-repeat/DRE sequences in place of the normal UAS (upstream activator sequence) of the *GAL1* promoter.

Figures 1A and 1B show how the yeast reporter strains were constructed. Figure 1A is a schematic diagram showing the screening



strategy. Yeast reporter strains were constructed that carried C-repeat/DRE sequences as UAS elements fused upstream of a *lacZ* reporter gene with a minimal *GAL1* promoter. The strains were transformed with an Arabidopsis expression library that contained random cDNA inserts fused to the GAL4  
5 activation domain (GAL4-ACT) and screened for blue colony formation on X-gal-treated filters. Figure 1B is a chart showing activity of the "positive" cDNA clones in yeast reporter strains. The oligonucleotides (oligos) used to make the UAS elements, and their number and direction of insertion, are indicated by the arrows.

10 Yeast strains carrying these reporter constructs produced low levels of beta-galactosidase and formed white colonies on filters containing X-gal. The reporter strains carrying the wild-type C-repeat/DRE sequences were transformed with a DNA expression library that contained random Arabidopsis cDNA inserts fused to the acidic activator domain of the yeast GAL4  
15 transcription factor, "GAL4-ACT" (Figure 1A). The notion was that some of the clones might contain a cDNA insert encoding a C-repeat/DRE binding domain fused to GAL4-ACT and that such a hybrid protein could potentially bind upstream of the *lacZ* reporter genes carrying the wild type C-repeat/DRE sequence, activate transcription of the *lacZ* gene and result in yeast forming  
20 blue colonies on X-gal-treated filters.

Upon screening about  $2 \times 10^6$  yeast transformants, three "positive" cDNA clones were isolated; i.e., clones that caused yeast strains carrying *lacZ* reporters fused to wild-type C-repeat/DRE inserts to form blue colonies on X-gal-treated filters (Figure 1B). The three cDNA clones did not cause a yeast  
25 strain carrying a mutant C-repeat/DRE fused to *LacZ* to turn blue (Figure 1B). Thus, activation of the reporter genes by the cDNA clones appeared to be dependent on the C-repeat/DRE sequence. Restriction enzyme analysis and DNA sequencing indicated that the three cDNA clones had an identical 1.8 kb insert (Figure 2A). One of the clones, designated pACT-11, was chosen for  
30 further study.

**D. Identification of 24 kDa polypeptide with an AP2 domain encoded by pACT-11.**

Figures 2A, 2B, 2C and 2D provide an analysis of the pACT-11 cDNA clone. Figure 2A is a schematic drawing of the pACT-11 cDNA insert indicating the location and 5' to 3' orientation of the 24 kDa polypeptide and 25s rRNA sequences. The cDNA insert was cloned into the *Xho*I site of the pACT vector. Figure 2B is a DNA and amino acid sequence of the 24 kDa polypeptide (SEQ ID NO:1 and SEQ ID NO:2). The AP2 domain is indicated by a double underline. The basic amino acids that potentially act as a nuclear localization signal are indicated with asterisks. The *Bcl*I site immediately upstream of the 24 kDa polypeptide used in subcloning the 24 kDa polypeptide and the *Eco*RV site used in subcloning the 3' end of CBF1 are indicated by single underlines. Figure 2C is a schematic drawing indicating the relative positions of the potential nuclear localization signal (NLS), the AP2 domain and the acidic region of the 24 kDa polypeptide. Numbers indicate amino acid residues. Figure 2D is a chart showing comparison of the AP2 domain of the 24 kDa polypeptide with that of the tobacco DNA binding protein EREBP2 (Okme-Takagi, M., et al., The Plant Cell 7:173-182 (1995) SEQ ID NOS: 10 and 11). Identical amino acids are indicated with single lines; similar amino acids are indicated by double dots; amino acids that are invariant in AP2 domains are indicated with asterisks (Klucher, K. M., et al., The Plant Cell 8:137-153 (1996)); and the histidine residues present in CBF1 and TINY (Wilson, K., et al., The Plant Cell 8:659-671 (1996)) that are tyrosine residues in all other described AP2 domains are indicated with a caret. A single amino acid gap in the CBF1 sequence is indicated by a single dot.

Our expectation was that the cDNA insert in pACT-11 would have a C-repeat/DRE binding domain fused to the yeast GAL4-ACT sequence. However, DNA sequence analysis indicated that an open reading frame of only nine amino acids had been added to the C-terminus of GAL4-ACT. It seemed highly unlikely that such a short amino acid sequence could comprise a DNA binding domain. Also surprising was the fact that about half of the cDNA insert in pACT-11 corresponded to 25s rRNA sequences (Figure 2A). Further analysis, however, indicated that the insert had an open reading frame, in opposite orientation to the GAL4-ACT sequence, deduced to encode a 24 kDa

polypeptide (Figure 2A-C). The polypeptide has a basic region that could potentially serve as a nuclear localization signal (Raikhel, N., *Plant Physiol.* **100**:1627-1632 (1992)) and an acidic C-terminal half (pI of 3.6) that could potentially act as an acidic transcription activator domain (Hahn, S., *Cell* **72**:481-483 (1993)). A search of the nucleic acid and protein sequence databases indicated that there was no previously described homology of the 24 kDa polypeptide. However, the polypeptide did have an AP2 domain (Jofuku, K. D., et al., *The Plant Cell* **6**:1211-1225 (1994)) (Figures 2B, D), a DNA binding motif of about 60 amino acids (Ohme-Takagi, M., et al., *The Plant Cell* **7**:173-182 (1994)) that is present in numerous plant proteins including the APETALA2 (Jofuku, K. D., et al., *The Plant Cell* **6**:1211-1225 (1994)), AINTEGUMENTA (Klucher, K. M., et al., *The Plant Cell* **8**:137-153 (1996); Elliot, R. C., et al., *The Plant Cell* **8**:155-168 (1996)) and TINY (Wilson, K., et al., *The Plant Cell* **8**:659-671 (1996)) proteins of Arabidopsis and the EREBPs (ethylene response element binding proteins) of tobacco (Ohme-Takagi, M., et al., *The Plant Cell* **7**:173-182 (1995)).

**E. 24 kDa polypeptide binds to the C-repeat/DRE and activates transcription in yeast.**

We hypothesized that the 24 kDa polypeptide was responsible for activating the *lacZ* reporter genes in yeast. To test this, the *BclI*-*BglII* fragment of pACT-11 containing the 24 kDa polypeptide, and the *BglII*-*BglII* fragment containing the 24 kDa polypeptide plus a small portion of the 25s rRNA sequence, was inserted into the yeast expression vector pDB20.1

Figure 3 is a chart showing activation of reporter genes by the 24 kDa polypeptide. Restriction fragments of pACT-11 carrying the 24 kDa polypeptide (*BclI*-*BglII*) or the 24 kDa polypeptide plus a small amount of 25s RNA sequence (*BglII*-*BglII*) were inserted in both orientations into the yeast expression vector pDB20.1 (see Figure 2A and 2B for location of *BclI* and *BglII* restriction sites). These "expression constructs" were transformed into yeast strains carrying the *lacZ* reporter gene fused to direct repeat dimers of either the wild-type *COR15a* C-repeat/DRE (oligonucleotide MT50) or the mutant M2*COR15a* C-repeat/DRE (oligonucleotide MT80). The specific activity of beta-galactosidase (nmoles o-nitrophenol produced/min<sup>-1</sup> x mg protein<sup>-1</sup>) was

determined from cultures grown in triplicate. Standard deviations are indicated. Abbreviations: pADC1, *ADC1* promoter; tADC1, *ADC1* terminator.

Plasmids containing either insert in the same orientation as the *ADC1* promoter stimulated synthesis of beta-galactosidase when transformed into yeast strains carrying the *lacZ* reporter gene fused to a wild-type *COR15a* C-repeat/DRE (Figure 3). The plasmids did not, however, stimulate synthesis of beta-galactosidase when transformed into yeast strains carrying *lacZ* fused to a mutant version of the *COR15a* C-repeat/DRE (Figure 3). These data indicated that the 24 kDa polypeptide could bind to the wild-type C-repeat/DRE and activate expression for the *lacZ* reporter gene in yeast. Additional experiments indicated that the 24 kDa polypeptide could activate expression of the *lacZ* reporter gene fused to either a wild-type *COR78* C-repeat/DRE (dimer of MT66) or a wild-type *COR15b* C-repeat/DRE (dimer of MT 68) (not shown). A plasmid containing the *BclI*-*BglII* fragment (which encodes only the 24 kDa polypeptide) cloned in opposite orientation to the *ADC1* promoter did not stimulate synthesis of beta-galactosidase in reporter strains carrying the wild-type *COR15a* C-repeat/DRE fused to *lacZ* (Figure 3). In contrast, a plasmid carrying the *BglII*-*BglII* fragment (containing the 24 kDa polypeptide plus some 25s rRNA sequences) cloned in opposite orientation to the *ADC1* promoter produced significant levels of beta-galactosidase in reporter strains carrying the wild-type *COR15a* C-repeat/DRE (Figure 3). Thus, a sequence located closely upstream of the 24 kDa polypeptide was able to serve as a cryptic promoter in yeast, a result that offered an explanation for how the 24 kDa polypeptide was expressed in the original pACT-11 clone.

**F. Gel shift analysis indicates that the 24 kDa polypeptide binds to the C-repeat/DRE.**

Gel shift experiments were conducted to demonstrate further that the 24 kDa polypeptide bound to the C-repeat/DRE. Specifically, the open reading frame for the 24 kDa polypeptide was inserted into the pET-28a(+) bacterial expression vector (see Materials and Methods) and the resulting 28 kDa fusion protein was expressed at high levels in *E. coli*. (Figure 4).

Figure 4 is a photograph of an electrophoresis gel showing expression of the recombinant 24 kDa polypeptide in *E. coli*. Shown are the results of

SDS-PAGE analysis of protein extracts prepared from *E. coli* harboring either the expression vector alone (vector) or the vector plus an insert encoding the 24 kDa polypeptide in sense (sense insert) or antisense (antisense insert) orientation. The 28 kDa fusion protein (see Materials and Methods) is indicated by an arrow.

Figure 5 is a photograph of a gel for shift assays indicating that CBF1 binds to the C-repeat/DRE. The C-repeat/DRE probe (1 ng) used in all reactions was a <sup>32</sup>P-labeled dimer of the oligonucleotide MT50 (wild type C-repeat/DRE from *COR15a*). The protein extracts used in the first four lanes were either bovine serum albumin (BSA) or the indicated CBF1 sense, antisense and vector extracts described in Figure 4. The eight lanes on the right side of the figure used the CBF1 sense protein extract plus the indicated competitor C-repeat/DRE sequences (100 ng). The numbers 1X, 2X and 3X indicate whether the oligonucleotides were monomers, dimers or trimers, respectively, of the indicated C-repeat/DRE sequences.

Protein extracts prepared from *E. coli* expressing the recombinant protein produced a gel shift when a wild-type *COR15a* C-repeat/DRE was used as probe (Figure 5). No shift was detected with BSA or *E. coli* extracts prepared from strains harboring the vector alone, or the vector with an antisense insert for the 24 kDa polypeptide. Oligonucleotides encoding wild-type C-repeat/DRE sequences from *COR15a* or *COR78* competed effectively for binding to the *COR15a* C-repeat/DRE probe, but mutant version of the *COR15a* C-repeat/DRE did not (Figure 5). These *in vitro* results corroborated the *in vivo* yeast expression studies indicating that the 24 kDa polypeptide binds to the C-repeat/DRE sequence. The 24 kDa polypeptide was thus designated CBF1 (C-repeat/DRE binding factor 1) and the gene encoding it named *CBF1*.

#### G. CBF1 is a unique or low copy number gene.

Figure 6 is a photograph of a southern blot analysis indicating *CBF1* is a unique or low copy number gene. Arabidopsis DNA (1 µg) was digested with the indicated restriction endonucleases and southern transfers were prepared and hybridized with a <sup>32</sup>P-labeled probe encoding the entire CBF1 polypeptide.

The hybridization patterns observed in southern analysis of Arabidopsis DNA using the entire *CBF1* gene as probe were relatively simple indicating that *CBF1* is either a unique or low copy number gene (Figure 6). The hybridization patterns obtained were not altered if only the 3' end of the gene was used as the probe (the *EcoRV/BglII* restriction fragment from pACT-11 encoding the acidic region of *CBF1*, but not the AP2 domain) or if hybridization was carried out at low stringency (not shown).

#### H. *CBF1* transcript level response to low temperature.

Figures 7A, 7B and 7C relate to *CBF1* transcripts in control and cold-treated Arabidopsis. Figure 7A is a photograph of a membrane RNA isolated from Arabidopsis plants that were grown at 22 °C or grown at 22 °C and transferred to 2.5 °C for the indicated times. Figures 7B and 7C are graphs showing relative transcript levels of *CBF1* and *COR15a* in control and cold-treated plants. The radioactivity present in the samples described in Figure 7A were quantified using a Betascope 603 blot analyzer and plotted as relative transcript levels (the values for the 22 °C grown plants being arbitrarily set as 1) after adjusting for differences in loading using the values obtained with the pHH25 probe.

Based on Figures 7A-7C, northern analysis indicated that the level of *CBF1* transcripts increased about 2 to 3 fold in response to low temperature (Figure 7B). In contrast, the transcript levels for *COR15a* increased approximately 35 fold in cold-treated plants (Figure 7C). Only a singly hybridizing band was observed for *CBF1* at either high or low stringency with probes for either the entire *CBF1* coding sequence or the 3' end of the gene (the *EcoRV/BglII* fragment of pACT-11) (not shown). The size of the *CBF1* transcripts was about 1.0 kb.

#### I. Discussion Of Experimental Results.

The above example regarding *CBF1* represents the first identification of a gene sequence which encodes a protein capable of binding to the C-repeat/DRE sequence CCGAC. The experimental results presented evidence

that CBF1 binds to the C-repeat/DRE both *in vitro* via gel shift assays and *in vivo* via yeast expression assays. Further, the results demonstrate that CBF1 can activate transcription of reporter genes in yeast that contain the C-repeat/DRE.

5           The results of the southern analysis indicate that *CBF1* is a unique or low copy number gene in *Arabidopsis*. However, the CBF1 protein contains a 60 amino acid motif, the AP2 domain, that is evolutionary conserved in plants (Weigel, D., *The plant Cell* 7:388-389 (1995)). It is present in the *APETALA2* (Jofuku, K. D., et al., *The Plant Cell* 6:1211-1225 (1994)), *AINTEGUMENTA* (Klucher, K. M., et al., *the Plant Cell* 8:137-153 (1996; and Elliot, R. C., et al., *The Plant Cell* 8:155-168 (1996)), *TINY* (Wilson, K., et al., *The Plant Cell* 8:659-671 (1996)) and cadmium-induced (Choi, S.-Y., et al., *Plant Physiol.* 108:849 (1995)) proteins of *Arabidopsis* and the EREBPs of tobacco (Ohme-Takagi, M. et al., *The Plant Cell* 7:173-182 (1995)). In addition, a search of the GenBank expressed sequence tagged cDNA database indicates that there is one cDNA from *B. napus*, two from *Ricinus communis*, and more than 25 from *Arabidopsis* and 15 from rice, that are deduced to encode proteins with AP2 domains. The results of Ohme-Takagi and Shinshi (Ohme-Takagi, M., et al., *The Plant Cell* 7:173-182 (1995)) indicate that the function of the AP2 domain is DNA-binding; this region of the putative tobacco transcription factor EREBP2 is responsible for its binding to the *cis*-acting ethylene response element referred to as the GCC-repeat. As discussed by Ohme-Takagi and Shinshi (Ohme-Takagi, M., et al., *the Plant Cell* 7:173-182 (1995)), the DNA-binding domain of EREBP2 (the AP2 domain) contains no significant amino acid sequence similarities or obvious structural similarities with other known transcription factors or DNA binding motifs. Thus, the domain appears to be a novel DNA-binding motif that to date, has only been found in plant proteins.

It is believed that the binding of CBF1 to the C-repeat/DRE involves the AP2 domain. In this regard, it is germane to note that the tobacco ethylene response element, AGCCGCC, closely resembles the C-repeat/DRE sequences present in the promoters of the *Arabidopsis* genes *COR15a*, *GGCCGAC*, and *COR78/RD29A*, *TACCGAC*. Applicants believe that CBF1, the EREBPs and other AP2 domain proteins are members of a superfamily of DNA binding proteins that recognize a family of *cis*-acting regulatory elements

having CCG as a common core sequence. Differences in the sequence surrounding the CCG core element could result in recruitment of different AP2 domain proteins which, in turn, could be integrated into signal transduction pathways activated by different environmental, hormonal and developmental cues. Such a scenario is akin to the situation that exists for the ACGT-family of *cis*-acting elements (Foster et al., FASEB J. 8:192-200 (1994)). In this case, differences in the sequence surrounding the ACGT core element result in the recruitment of different bZIP transcription factors involved in activating transcription in response to a variety of environmental and developmental signals.

The results of the yeast transformation experiments indicate that CBF1 has a domain that can serve as a transcriptional activator. The most likely candidate for this domain is the acidic C-terminal half of the polypeptide. Indeed, random acidic amino acid peptides from *E. coli* have been shown to substitute for the GAL4 acidic activator domain of GAL4 in yeast (Ma, J. and M. Ptashne, Cell 51:113-199 (1987)). Moreover, acidic activator domains have been found to function across kingdoms (Hahn, S., Cell 72:481-483 (1993)); the yeast GAL4 acidic activator, for instance, can activate transcription in tobacco (Ma, J., et al., Nature 334:631-633 (1988)). It has also been shown that certain plant transcription factors, such as Vp1 (McCarty, D. R., et al., Cell 66:895-905 (1991)), have acidic domains that function as transcriptional activators in plants. Significantly, the acidic activation domains of the yeast transcription factors VP16 and GCN4 require the "adaptor" proteins ADA2, ADA3, and GCN5 for full activity (see Guarente, L., Trends Biochem. Sci. 20:517-521 (1995)). These proteins form a heteromeric complex (Horiuchi, J., et al., Mol. Cell Biol. 15:1203-1209 (1995)) that bind to the relevant activation domains. The precise mechanism of transcriptional activation is not known, but appears to involve histone acetylation: there is a wealth of evidence showing a positive correlation between histone acetylation and the transcriptional activity of chromatin (Wolffe, A. P., Trends Biochem. Sci. 19:240-244 (1994)) and recently, the GCN5 protein has been shown to have histone acetyltransferase activity (Brownell, J. E., et al., Cell 84:843-851 (1996)). Genetic studies indicate that CBF1, like VP16 and GCN4, requires ADA2, ADA3 and GCN5 to function optimally in yeast. The fundamental question thus raised is whether



plants have homologs of ADA2, ADA3 and GCN5 and whether these adaptors are required for CBF1 function (and function of other transcription factors with acidic activator regions) in Arabidopsis.

5 A final point regards regulation of CBF1 activity. The results of the northern analysis indicate that *CBF1* transcript levels increase only slightly in response to low temperature, while those for *COR15a* increase dramatically (Fig. 7). Thus, unlike in yeast, it would appear that transcription of *CBF1* in Arabidopsis at warm temperatures is not sufficient to cause appreciable activation of promoters containing the C-repeat/DRE. The molecular basis for 10 this apparent low temperature activation of CBF1 in Arabidopsis is not known. One intriguing possibility, however is that CBF1 might be modified at low temperature in Arabidopsis resulting in either stabilization of the protein, translocation of the protein from the cytoplasm to the nucleus, or activation of either the DNA binding domain or activation domain of the protein. Such 15 modification could involve a signal transduction pathway that is activated by low temperature. Indeed, as already discussed, cold-regulated expression of COR genes in Arabidopsis and alfalfa appears to involve a signal transduction pathway that is activated by low temperature-induced calcium flux (Knight, H., et al., The Plant Cell 8:489-503 (1996); Knight, M. R., et al., Nature 352:524-526 (1991); Monroy, A. F., et al, Plant Physiol. 102:1227-1235 (1993); Monroy, 20 A. F., and R. S., The Plant Cell, 7:321-331 (1995)). It will, therefore, be of interest to determine whether CBF1 is modified at low temperature, perhaps by phosphorylation, and if so, whether this is dependent on calcium-activated signal transduction.

25

## 2. Use of CBF1 To Induce Cold Regulated Gene Expression in Nonacclimated *Arabidopsis* Plants.

The following example demonstrates that increased expression of 30 CBF1 induces *COR* gene expression in nonacclimated *Arabidopsis* plants. Transgenic Arabidopsis plants that overexpress *CBF1* were created by placing a cDNA encoding CBF1 under the control of the strong cauliflower mosaic virus (CaMV) 35S promoter and transforming the chimeric gene into *Arabidopsis* ecotype RLD plants (Standard procedures were used for plasmid 35 manipulations (J. Sambrook, et al., Molecular Cloning, A Laboratory Manual

(Cold Spring Harbor Laboratory Press, Cold Spring Harbor, ed. 2, (1989)). The CBF1-containing *Asel-Bgl* fragment from pACT-Bgl + (Stockinger, E. J., et al., Proc. Natl. Acad. Sci. U.S.A. **94**:1035 (1997)) was gel-purified, *Bam*HI linkers were ligated to both ends and the fragment was inserted into the *Bam*HI site in pCIB710 (S. Rothstein, et al., Gene **53**:153-161 (1987)) which contains the CaMV 35S promoter and terminator. The chimeric plasmid was linearized at the *Kpn*I site and inserted into the *Kpn*I site of the binary vector pCIB10g (Ciba-Geigy, Research Triangle Park, NC). The plasmid was transformed into *Agrobacterium tumefaciens* strain C58C1 (pMP90) by electroporation. *Arabidopsis* plants were transformed by the vacuum infiltration procedure (N. Bechtold, J. Ellis, and G. Pelletier, C. R. Acad. Sci. Paris, Life Sci. **316**:1194-1199 (1993)) as modified (A. van Hoof, P. J. Green, Plant Journal **10**:415-424 (1996)). Initial screening gave rise to two transgenic lines, A6 and B16, that accumulated CBF1 transcripts at elevated levels.

Figure 8 is a Northern blot showing CBF1 and COR transcript levels in RLD and transgenic *Arabidopsis* plants. Leaves from nonacclimated and three-day cold-acclimated plants (*Arabidopsis thaliana* ecotype RLD plants were grown in pots under continuous light (100  $\mu$ E/m<sup>2</sup>/sec) at 22 C for 18-25 days as described (Gilmour, S. J., et al., Plant Physiol. **87**:735 (1988)). In some cases, plants were then cold-acclimated by placing them at 2.5 °C under continuous light (50  $\mu$ E/m<sup>2</sup>/sec) for varying amounts of time. Leaves were harvested and total RNA prepared and analyzed for CBF1 and COR transcripts by RNA blot analysis using <sup>32</sup>P-radiolabeled probes (Total RNA was isolated from plant leaves and subjected to RNA blot analysis using high stringency hybridization and wash conditions as described (E.J. Stockinger, et al., Proc. Natl. Acad. Sci. USA **94**:1035 (1997); and S.J. Gilmour, et al., Plant Physiol. **87**:735 (1988)).

Figure 9 is an immunoblot showing COR15am protein levels in RLD and transgenic *Arabidopsis* plants. Total soluble protein (100  $\mu$ g) was prepared from leaves of the nonacclimated RLD (RLDw), 4-day cold-acclimated RLD (RLDc4), 7-day cold-acclimated RLD (RLDc7) and nonacclimated A6 and B16 plants and the levels of COR15am determined by immunoblot analysis using antiserum raised against the COR15am polypeptide (Total soluble protein was isolated from plant leaves, fractionated by tricine

SDS-PAGE and transferred to 0.2 micron nitrocellulose as previously described (N. N. Artus et al., Proc. Natl. Acad. Sci. U.S.A. **93**:13404 (1996)). COR15am protein was detected using antiserum raised to purified COR15am and protein A conjugated alkaline phosphatase (Sigma, St. Louis, MO) (N. N. Artus et al., Proc. Natl. Acad. Sci. U.S.A. **93**:13404 (1996)). No reacting bands were observed with preimmune serum (not shown).

Southern analysis indicated that the A6 line had a single DNA insert while the B16 line had multiple inserts (not shown). Examination of fourth generation homozygous A6 and B16 plants indicated that *CBF1* transcript levels were higher in nonacclimated A6 and B16 plants than they were in nonacclimated RLD plants, the levels in A6 being about three fold higher than in B16 (Figure 8).

*CBF1* overexpression resulted in strong induction of *COR* gene expression (Figure 8). Specifically, the transcript levels of *COR6.6*, *COR15a*, *COR47* and *COR78* were dramatically elevated in nonacclimated A6 and B16 plants as compared to nonacclimated RLD plants. The effect was greater in the A6 line, where *COR* transcript levels in nonacclimated plants approximated those found in cold-acclimated RLD plants. The finding that *COR* gene expression was greater in A6 plants than in B16 plants was consistent with *CBF1* transcript levels being higher in the A6 plants (Figure 7A). Immunoblot analysis indicated that the levels of the COR15am (Figure 9) and COR6.6 (not shown) polypeptides were also elevated in the A6 and B16 lines, the level of expression again being higher in the A6 line. Attempts to identify the *CBF1* protein in either RLD or transgenic plants were unsuccessful. Overexpression of *CBF1* had no effect on the transcript levels for *eIF4A* (eukaryotic initiation factor 4A) (Metz, A.M., et al., Gene **120**:313 (1992)), a constitutively expressed gene that is not responsive to low temperature (Figure 8) and had no obvious effects on plant growth and development.

The results from this example demonstrate that overexpression of the *Arabidopsis* transcriptional activator *CBF1* induces expression of an *Arabidopsis* *COR* "regulon" composed of genes carrying the CRT/DRE DNA regulatory element. It appears that *CBF1* binds to the CRT/DRE DNA regulatory elements present in the promoters of these genes and activates transcription which is consistent with the notion of *CBF1* having a role in *COR*

gene regulation. Significantly, there was a strong correlation between CBF1 transcript levels and the magnitude of *COR* gene induction in nonacclimated A6, B16, and RLD plants (Figure 8). However, upon low temperature treatment the level of CBF1 transcripts remained relatively low in RLD plants, while *COR* gene expression was induced to about the same level as that in nonacclimated A6 plants (Figure 8). Thus, it appears that CBF1 or an associated protein becomes "activated" in response to low temperature.

### 3. ***CBF1* Overexpression Resulted in a Marked Increase in Plant Freezing Tolerance**

The following example describes a comparison of the freezing tolerance of nonacclimated *Arabidopsis* plants which overexpress CBF1 to that of cold-acclimated wild-type plants. As described below, the freezing tolerance of nonacclimated *Arabidopsis* plants overexpressing CBF1 significantly exceeded that of non-acclimated wild-type *Arabidopsis* plants and approached that of cold-acclimated wild-type plants.

Freezing tolerance was determined using the electrolyte leakage test (Sukumaran, N. P., et al., HortScience 7:467 (1972)). Detached leaves were frozen to various subzero temperatures and, after thawing, cellular damage (due to freeze-induced membrane lesions) was estimated by measuring ion leakage from the tissues.

Figures 10A and 10B are graphs showing freezing tolerance of leaves from RLD and transgenic *Arabidopsis* plants. Leaves from nonacclimated RLD (RLDw) plants, cold-acclimated RLD (RLDc) plants and nonacclimated A6, B16 and T8 plants were frozen at the indicated temperatures and the extent of cellular damage was estimated by measuring electrolyte leakage (Electrolyte leakage tests were conducted as described (N.P.Sukumaran, et al., HortScience 7, 467 (1972); and S.J. Gilmour, et al., Plant Physiol. 87:735 (1988)) with the following modifications. Detached leaves (2-4) from nonacclimated or cold-acclimated plants were placed in a test tube and submerged for 1 hour in a -2 °C water-ethylene glycol bath in a completely randomized design, after which ice crystals were added to nucleate freezing. After an additional hour of incubation at -2 °C, the samples were cooled in decrements of 1 °C each hour until -8 °C was reached. Samples (five

replicates for each data point) were thawed overnight on ice and incubated in 3 ml distilled water with shaking at room temperature for 3 hours. Electrolyte leakage from leaves was measured with a conductivity meter. The solution was then removed, the leaves frozen at -80 °C (for at least one hour), and the solution returned to each tube and incubated for 3 hours to obtain a value for 100% electrolyte leakage. In Figures 10A and 10B, the RLDc plants were cold-acclimated for 10 and 11 days, respectively. Error bars indicate standard deviations.

As can be seen from Figure 10A and 10B, *CBF1* overexpression resulted in a marked increase in plant freezing tolerance. The experiment presented in Figure 10A indicates that the leaves from both nonacclimated A6 and B16 plants were more freezing tolerant than those from nonacclimated RLD plants. Indeed, the freezing tolerance of leaves from nonacclimated A6 plants approached that of leaves from cold-acclimated RLD plants. The results also indicate that the leaves from nonacclimated A6 plants were more freezing tolerant than those from nonacclimated B16 plants, a result that is consistent with the greater level of *CBF1* and *COR* gene expression in the A6 line.

The results presented in Figure 10B further demonstrate that the freezing tolerance of leaves from nonacclimated A6 plants was greater than that of leaves from nonacclimated RLD plants and that it approached the freezing tolerance of leaves from cold-acclimated RLD plants. In addition, the results indicate that overexpression of *CBF1* increases freezing tolerance to a much greater extent than overexpressing *COR15a* alone. This conclusion comes from comparing the freezing tolerance of leaves from nonacclimated A6 and T8 plants (Figure 10B). T8 plants (Artus, N. N., et al., Proc. Natl. Acad. Sci. U.S.A. 93:13404 (1996)) are from a transgenic line that constitutively expresses *COR15a* (under control of the CaMV 35S promoter) at about the same level as in A6 plants (Figure 1). However, unlike in A6 plants, other CRT/DRE-regulated *COR* genes are not constitutively expressed in T8 plants (Figure 8).

A comparison of  $EL_{50}$  values (the freezing temperature that results in release of 50% of tissue electrolytes) of leaves from RLD, A6, B16 and T8 plants is presented in Table 2.

EL<sub>50</sub> values were calculated and compared by analysis of variance curves fitting up to third order linear polynomial trends were determined for each electrolyte leakage experiment. To insure unbiased predictions of electrolyte leakage, trends significantly improving the model fit at the 0.2 probability level were retained. EL<sub>50</sub> values were calculated from the fitted models. In Table 2, an unbalanced one-way analysis of variance, adjusted for the different numbers of EL<sub>50</sub> values for each plant type, was determined using SAS PROC GLM [SAS Institute, Inc. (1989), SAS/STAT User's Guide, Version 6, Cary, NC]. EL<sub>50</sub> values  $\pm$  SE (n) are presented on the diagonal line for leaves from nonacclimated RLD (RLDw), cold-acclimated (7 to 10 days) RLD (RLDc) and nonacclimated A6, B16 and T8 plants. P values for comparisons of EL<sub>50</sub> values are indicated in the intersecting cells.

TABLE 2

EL <sub>50</sub> values					
	RLDw	RLDc	A6	B16	T8
RLDw	-3.9 $\pm$ 0.21 (8)	P<0.0001	P<0.0001	P=0.0014	P=0.7406
RLDc		-7.6 $\pm$ 0.30 (4)	P=0.3261	P<0.0001	P<0.0001
A6			-7.2 $\pm$ 0.25 (6)	P<0.0001	P<0.0001
B16				-5.2 $\pm$ 0.27 (5)	P=0.0044
T8					-3.8 $\pm$ 0.35 (3)

The data confirm that: 1) the freezing tolerance of leaves from both nonacclimated A6 and B16 plants is greater than that of leaves from both nonacclimated RLD and T8 plants; and 2) that leaves from nonacclimated A6 plants are more freezing tolerant than leaves from nonacclimated B16 plants. No significant difference was detected in EL<sub>50</sub> values for leaves from

nonacclimated A6 and cold-acclimated RLD plants or from nonacclimated RLD and T8 plants.

5 The enhancement of freezing tolerance in the A6 line was also apparent at the whole plant level. Figure 11 is a photograph showing freezing survival of RLD and A6 *Arabidopsis* plants. Nonacclimated (WARM) RLD and A6 plants and 5-day cold-acclimated (COLD) RLD plants were frozen at -5 °C for 2 days and then returned to a growth chamber at 22 °C (Pots (3.5 inch) containing about 40 nonacclimated *Arabidopsis* plants (20 day old) and 4 day cold-acclimated plants (25 days old) (*Arabidopsis thaliana* ecotype RLD plants were grown in pots under continuous light ( 100  $\mu\text{E}/\text{m}^2/\text{sec}$ ) at 22 °C for 18-25 days as described (S.J. Gilmour, et al., Plant Physiol. **87**:735 (1988)). In some cases, plants were then cold-acclimated by placing them at 2.5 °C under continuous light ( 50  $\mu\text{E}/\text{m}^2/\text{sec}$ ) for varying amounts of time) were placed in a completely randomized design in a -5 °C cold chamber in the dark. After 1 hour, ice chips were added to each pot to nucleate freezing. Plants were removed after 2 days and returned to a growth chamber at 22 °C.). A photograph of the plants after 7 days of regrowth is shown.

15 Although the magnitude of the difference varied from experiment to experiment, nonacclimated A6 plants consistently displayed greater freezing tolerance in whole plant freeze tests than did nonacclimated RLD plants (Figure 11). No difference in whole plant freeze survival was detected between nonacclimated B16 and RLD plants or nonacclimated T8 and RLD plants (not shown).

20 The results of this experiment show that CBF1-induced expression of CRT/DRE-regulated *COR* genes result in a dramatic increase in freezing tolerance and confirms the belief that *COR* genes play a major role in plant cold acclimation. The increase in freezing tolerance brought about by expressing the battery of CRT/DRE-regulated *COR* genes was much greater than that brought about by overexpressing *COR15a* alone indicating that *COR* genes in addition to *COR15a* have roles in freezing tolerance.

25 Traditional plant breeding approaches have met with limited success in improving the freezing tolerance of agronomic plants (Thomashow, M. F., Adv. Genet **28**:99 (1990)). For instance, the freezing tolerance of the best wheat varieties today is essentially the same as the most freezing-tolerance varieties

developed in the early part of this century. Thus, in recent years there has been considerable interest that biotechnology might offer new strategies to improve the freezing tolerance of agronomic plants. By the results of the present invention, Applicants demonstrate the ability to enhance the freezing tolerance of nonacclimated *Arabidopsis* plants by increasing the expressing of the *Arabidopsis* regulatory gene *CBF1*. As described throughout this application, the ability of the present invention to modify the expression of environmental stress tolerance genes such as core genes has wide ranging implications since the CRT/DRE DNA regulatory element is not limited to *Arabidopsis* (Jiang C., et al., Plant Mol. Biol. **30**:679 (1996)). Rather, *CBF1* and homologous genes can be used to manipulate expression of CRT/DRE-regulated *COR* genes in important crop species and thereby improve their freezing tolerance. By transforming modified versions of *CBF1* (or homologs) into such plants, it will extend their safe growing season, increase yield and expand areas of production.

#### 4. Selection Of Promoters To Control Expression of *CBF1* in Plants

The following examples describe the isolation of different promoters from plant genomic DNA, construction of the plasmid vectors carrying the *CBF1* gene and the inducible promoters, transformation of *Arabidopsis* cells/plants with these constructs, and regeneration of transgenic plants with increased tolerance to environmental stresses.

##### A. Isolation of inducible promoters from plant genomic DNAs

Inducible promoters from different plant genomic DNAs were identified and isolated by PCR amplification using primers designed to flank the promoter region and contain suitable restriction sites for cloning into the expression vector. The following genes were used to BLAST search Genbank to find the inducible promoters: *Dreb2a*; *P5CS*; *Rd22*; *Rd29a*; *Rd29b*; *Rab18*; *Cor47*. Table 3 lists the accession numbers and positions of these promoters. Table 4 lists the forward and reverse primers that were used to isolate the promoters.



TABLE 3

Gene Name	Accession No.	Position	Length (bps)
Dreb2a	AB010692	51901-53955	2054
P5CS	AC003000	45472-47460	1988
Rd22	D10703	17-1046	1029
Rd29a	D13044	3870-5511	1641
Rd29b	D13044	90-1785	1695
Rab18	AB013389	8070-9757	1687
Cor47	AB004872	1-1370	1370

TABLE 4

5

Promoter name	Primer name	Cloning sites	SEQ.ID.No.
Dreb2a	Dreb2a-reverse	HindIII (AAGCTT)	19
	Dreb2a-forward	BglII (AGATCT)	20
P5CS	P5CS-reverse	HindIII (AAGCTT)	21
	P5CS-forward	BglII (AGATCT)	22
Rd22	Rd22-reverse	HindIII (AAGCTT)	23
	Rd22-forward	KpnI (GGTACC)	24
Rd29a	Rd29a-reverse	HindIII (AAGCTT)	25
	Rd29a-forward	KpnI (GGTACC)	26
Rd29b	Rd29b-reverse	HindIII (AAGCTT)	27
	Rd29b-forward	KpnI (GGTACC)	28
Rab18	Rab18-reverse	HindIII (AAGCTT)	29
	Rab18-forward	BglII (AGATCT)	30
Cor47	Cor47-reverse	HindIII (AAGCTT)	31
	Cor47-forward	BglII (AGATCT)	32

(1) *Dreb2a promoter*

10 A cDNA encoding DRE (C-repeat) binding protein (DREB2A) has been recently identified (Liu, et al. 1998 Plant Cell 10:1391-1406). The transcription of the DREB2A gene is activated by dehydration and high-salt stress, but not by cold stress. The upstream untranslated region (166 bps) of *dreb2a* was used to BLAST-search the public database. A region containing the DREB2A

promoter was identified in chromosome 5 of *Arabidopsis* (Accession No. AB010692) between nucleotide positions 51901-53955 (Table 3).

Two PCR primers designed to amplify the promoter region from *Arabidopsis thaliana* genomic DNA are as follows: *dre2a*-reverse:

5 5'-GCCCAAGCTTCAAGTTTAGTGAGCACTATGTGCTCG-3' [SEQ ID No. 19];  
and *dre2a*-forward: 5'-GGAAGATCTCCTTCCCAGAAACAACACAATCTAC-3'  
[SEQ. ID. No. 20]. The *dre2a*-reverse primer includes a Hind III (AAGCTT) restriction site near the 5'-end of the primer and *dre2a*-forward primer has a Bgl II (AGATCT) restriction site at near 5'-end of the primer. These restriction  
10 sites may be used to facilitate cloning of the fragment into an expression vector.

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the  
15 genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The reaction conditions that may be used in this PCR experiment are as follows: Segment 1: 94°C, 2 minutes; Segment 2: 94°C, 30 seconds; 60°C, 1 minute; 72°C, 3 minutes, for a total of 35 cycles; Segment 3: 72°C for 10 minutes. A PCR  
20 product of 2054 bp is expected.

The PCR products can be subject to electrophoresis in a 0.8% agarose gel and visualized by ethidium bromide staining. The DNA fragments containing the inducible promoter will be excised and purified using a Qiaquick gel extraction kit (Qiagen, CA).

## 25 (2) *P5CS promoter*

A cDNA for delta 1-pyrroline-5-carboxylate synthetase (P5CS) has been isolated and characterized (Yoshida, et al., 1995, Plant J. 7:751-760). The cDNA encodes an enzyme involved in the biosynthesis of proline under  
30 osmotic stress (drought/high salinity). The transcription of the P5CS gene was found to be induced by dehydration, high salt and treatment with plant hormone ABA, while it did not respond to heat or cold treatment.

A genomic DNA containing a promoter region of P5CS was identified by a BLAST search of Genbank using the upstream untranslated region (106 bps)

of the *P5CS* sequence (Accession No. D32138). The sequence for the *P5CS* promoter is located in the region between from nucleotide positions 45472 to 47460 (Accession No. AC003000; Table 3).

Reverse and forward PCR primers designed to amplify this promoter region from *Arabidopsis thaliana* genomic DNA are *P5CS*-reverse primer 5'-GCCCAAGCTTGTTCATTTCTCCATGAAGGAGAT-3' [SEQ. ID. No. 21]; and *P5CS*-forward primer 5'-GGAAGATCTTATCGTCGTCGTCGTCTACCAAACCACAC-3' [SEQ. ID. No. 22].

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1988 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

### (3) *rd22* promoter

A cDNA clone of *rd22* was isolated from *Arabidopsis* under dehydration conditions (Yamaguchi-Shinozaki and Shinozaki, Mol. Gen. Genet. **238**:17-25 (1993)). Transcripts of *rd22* were found to be induced by salt stress, water deficit and endogenous abscisic acid (ABA) but not by cold or heat stress. A promoter region was identified from Genebank by using Nucleotide Search WWW Entrez at the NCBI with the *rd22* as a search word. The sequence for the *rd22* promoter is located in the region between nucleotide positions 17 to 1046 (Accession No. D10703; Table 3).

Reverse and forward PCR primers designed to amplify this promoter region from *Arabidopsis thaliana* genomic DNA are *rd22*-reverse primer 5'-GCTCTAAGCTTCACAAGGGGTTTCGTTTGGTGC-3' [SEQ. ID. No. 23]; and *rd22*-forward primer 5'-GGGGTACCTTTTGGGAGTTGGAATAGAAATGGGTTTGATG-3' [SEQ. ID. No. 24]. The *rd22*-reverse primer includes a Hind III (AAGCTT) restriction site near the 5'-end of primer and *rd22*-forward primer has a KpnI (GGTACC) restriction site at near 5'-end of primer.

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1029 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

(4) *rd29a* promoter

The *rd29a* and *rb29b* genes were isolated and characterized by Shinozaki's group in Japan (Yamaguchi-Shinozaki and Shinozaki, Plant Physiol. **101**: 1119-1120 (1993)). Both *rd29a* and *rb29b* gene expressions were found to be induced by desiccation, salt stress and exogenous ABA treatment (Yamaguchi-Shinozaki and Shinozaki, Plant Physiol. **101**: 1119-1120 (1993); Ishitani et al., Plant Cell **10**: 1151-1161 (1998)). The *rd29a* gene expression was induced within 20 min after desiccation, but *rd29b* mRNA did not accumulate to a detectable level until 3 hours after desiccation. Expression of *rd29a* could also be induced by cold stress, whereas expression of *rd29b* could not be induced by low temperature.

A genomic clone carrying the *rd29a* promoter was identified by using Nucleotide Search WWW Entrez at the NCBI with the *rd29a* as a search word. The sequence for the *rd29a* promoter is located in the region between nucleotide positions 3870 to 5511 (Accession No. D13044, Table 3).

Reverse and forward primers designed to amplify this promoter region from *Arabidopsis* genomic DNA are: *rd29a*-reverse primer 5'-GCCCAAGCTTAATTTTACTCAAAATGTTTTGGTTGC-3' [SEQ. ID. No. 25]; and *rd29a*-forward primer 5'-CCGGTACCTTTCCAAAGATTTTTTCTTTCCAATAGAAGTAATC-3' [SEQ. ID. No. 26]. The *rd29a*-reverse primer includes a Hind III (AAGCTT) restriction site near the 5'-end of primer and *rd29a*-forward primer has a KpnI (GGTACC) restriction site near 5'-end of primer.

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the

genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1641 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

5

(5) *rd29b* promoter

A genomic clone carrying the *rd29b* promoter was identified by using Nucleotide Search WWW Entrez at the NCBI with the *rd29b* as a search word. The sequence for the *rd29a* promoter was located in the region between  
10 nucleotide positions 90 to 1785 for *rd29b* (Accession No. D13044; Table 3).

Reverse and forward PCR primers designed to amplify this promoter region from *Arabidopsis thaliana* genomic DNA are: *rd29b*-reverse primer 5'-GCGGAAGCTTCATTTTCTGCTACAGAAGTG-3' [SEQ. ID. No. 27]; and  
15 *rd29b*-forward primer 5'-CCGGTACCTTTCCAAAGCTGTGTTTTCTCTTTTCAAGTG-3' [SEQ. ID. No. 28].

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the  
20 genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1695 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

25 (6) *rab18* promoter

A *rab*-related (responsive to ABA) gene, *rab18* from *arabidopsis* has been isolated. The gene encodes a hydrophilic, glycine-rich protein with the conserved serine- and lysine-rich domains. The *rab18* transcripts accumulate in plants exposed to water deficit or exogenous abscisic acid (ABA) treatment.  
30 A weak induction of *rab18* mRNA by low temperature was also observed (Ishitani et al., Plant Cell 10: 1151-1161 (1998)).

A genomic DNA containing a promoter region of *rab18* was identified by a BLAST search of Genbank using the upstream untranslated region (757 bps) of the *rab18* sequence (Accession No. L04173). The sequence of the *rab18*

promoter is located in the region between nucleotide positions 8070 to 9757 (Accession No. AB013389).

Reverse and forward PCR primers designed and used to amplify this promoter region from *Arabidopsis thaliana* genomic DNA are: *rab18*-reverse primer

5'-GCCCAAGCTTCAAATTCTGAATATTCACATATCAAAAAAGTG-3' [SEQ. ID. No. 29]; and *rab18*-forward primer 5'-

GGAAGATCTGTTCTTCTTGTCTTAAGCAAACACTTTGAGC-3' [SEQ. ID. No. 30]. The *rab18*-reverse primer includes a Hind III (AAGCTT) restriction site near the 5'-end of the primer and *rab18*-forward primer has a Bgl II (AGATCT) restriction site near the 5'-end of the primer.

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1687 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

#### (7) *Cor47* promoter

The DNA sequence of cDNA for cold-regulated (*cor47*) gene of *Arabidopsis thaliana* was determined. Gilmour et al., Plant Molecular Biology 18: 13-21 (1992)). Expression of *cor47* gene was induced by cold stress, dehydration and high NaCl treatment (Ishitani et al., Plant Cell, 10: 1151-1161 (1998)). The promoter region of *cor47* gene was identified in Genbank by using Nucleotide Search WWW Entrez at the NCBI with the *cor47* as a search word. The sequence of the *cor47* promoter is located in the region between nucleotide positions 1-1370 (Accession No. AB004872; Table 3).

Reverse and forward PCR primers designed to amplify this promoter region from *Arabidopsis thaliana* genomic DNA are: *cor47*-reverse primer 5'-GCCCAAGCTTTCGTCTGTTATCATACAAGGCACAAAACGAC-3' [SEQ. ID. No. 31]; and *cor47*-forward primer 5'-GGAAGATCTAGTTAATCTTGATTTGATTAAAGTTTATATAG-3' [SEQ. ID. No. 32]. The *cor47*-reverse primer includes a Hind III (AAGCTT) restriction

site near the 5'-end of the primer and *cor47*-forward primer has a Bgl II (AGATCT) restriction site near the 5'-end of the primer.

Total genomic DNA may be isolated from *Arabidopsis thaliana* (ecotype colombia) by using the CTAB method (Ausubel et al. (1992) Current Protocols in Molecular Biology (Greene & Wiley, New York)). Ten nanograms of the genomic DNA can be used as a template in a PCR reaction under conditions suggested by the manufacturer (Boehringer Mannheim). The PCR product is expected to be 1370 bps and may be PCR amplified and gel purified following the same protocol described for the *dreb2a* promoter.

#### **B. Isolation of seed-specific promoters from plant genomic DNAs**

The napin promoter from *Brassica campestris* (genbank accession number M64632) is a seed-specific promoter. A fragment of the napin promoter (between nucleotides 1146 to 2148) is identified and isolated by PCR amplification using a 5' PCR primer containing a HindIII site upstream of the promoter and a 3' PCR primer containing a BamHI site downstream of the promoter. Deletions of the napin promoter to -211 and -152 have been shown to have reduced levels of expression (Ellerstrom et al. Plant Mol Biol **32**:1019-27 (1996); Stalberg et al. Planta **199**: 515-9 (1996); Stalberg et al. Plant Mol Biol **23**: 671-83 (1993)). These 5' deleted promoters are useful to have reduced levels of CBF1 expression for applications where the larger napin promoter fragment is too strong.

Other seed-active promoters or deletions of these promoters can also be isolated from genomic DNA by using the same method described above for the napin promoter. Examples of these promoters include but are not limited to the soybean 7S seed storage protein (Chen et al., Developmental Genetics **10**:112-122 (1989), the bean phaseolin promoter (cited in US Patent No. 5,003,045), the Arabidopsis 12S globulin (cruiferin) promoter (Pang, et al., Plant Mol. Biol. **11**:805-820 (1988), the maize globulin1 promoter (Kriz et al. Plant Physiol. **91**:636 (1989); US Patent No. 5,773,691). These promoters may be used for altering COR gene expression in cereals such as corn, barley, wheat, rice and rye seeds.

**C. Construction of the plasmids containing CBF1 and inducible or tissue-specific promoter**

5 The expression binary vector pMEN020 contains a kanamycin resistance gene (neomycin phosphotransferase) for antibiotic selection of the transgenic plants and a *Spc/Str* gene used for bacterial or agrobacterial selections. The pMEN020 plasmid is digested with restriction enzymes such as HindIII and BglII to remove the 35S promoter. The 35S promoter is then replaced with an inducible promoter.

10 The expression binary vector pMEN050 is derived from pMEN020 by replacing the NptII kanamycin resistance gene with the Bar gene (US Patent 5,646, 024). PMEN050 is digested with HindIII and BamHI restriction enzymes to remove the EcaMV 35S promoter. The EcaMV 35S promoter is then replaced with the seed-specific napin promoter, resulting plasmid pMEN1001.  
15 Similarly, the EcaMV 35S promoter is also replaced with the seed-specific napin promoters with -211 and -152 end point deletions to generate plasmid pMEN1002 and pMEN1003, respectively.

(1) *Cloning of the inducible promoter into pMEN020*

20 The sequences of the inducible promoters that are PCR amplified and gel purified, as well as the plasmid pMEN020, are subject to restriction digestion with their respective restriction enzymes as listed in Table 4. Both DNA samples are purified by using the Qiaquick purification kit (Qiagen, CA) and ligated at a ratio of 3:1 (vector to insert). Ligation reactions using T4 DNA ligase (New England Biolabs, MA) are carried out at 16°C for 16 hours. The  
25 ligated DNAs are transformed into competent cells of the *E. coli* strain DH5 by using the heat shock method. The transformed cells are plated on LB plates containing 100 µg/ml spectinomycin (Sigma). Individual colonies are grown overnight in five milliliters of LB broth containing 100 µg/ml spectinomycin at 37°C.

30 Plasmid DNAs from transformants are purified by using Qiaquick Mini Prep kits (Qiagen, CA) according to the manufacturer's instruction. The presence of the promoter insert is verified by restriction mapping with the respective restriction enzymes as listed in Table 4 to cut out the cloned insert. The plasmid DNA is also subject to double-strand DNA sequencing analysis



using a vector primer (E9.1 primer 5'-CAAACCTCAGTAGGATTCTGGTGTGT-3'  
[SEQ. ID. No. 33].

(2) *Cloning of the cbf1 gene into the plasmids containing the inducible  
promoters*

To clone the CBF1 gene into the plasmids, different PCR primers with  
suitable restriction sites for each plasmid are used to isolate *cbf1* gene from  
*Arabidopsis thaliana* genomic DNA. The primers that may be used are listed in  
Table 5.

**TABLE 5**

Promoter name	Primer name	Cloning sites
Dreb2a	Cbf1-reverse1	BglII (AGATCT)
	Cbf1-forward1	BamHI (GGATCC)
P5CS	Cbf1-reverse1	BglII (AGATCT)
	Cbf1-forward1	BamHI (GGATCC)
Rd22	Cbf1-reverse2	KpnI (GGTACC)
	Cbf1-forward1	BamHI (GGATCC)
Rd29a	Cbf1-reverse2	KpnI (GGTACC)
	Cbf1-forward1	BamHI (GGATCC)
Rd29b	Cbf1-reverse2	KpnI (GGTACC)
	Cbf1-forward1	BamHI (GGATCC)
Rab18	Cbf1-reverse1	BglII (AGATCT)
	Cbf1-forward2	XbaI (TCTAGA)
Cor47	Cbf1-reverse1	BglII (AGATCT)
	Cbf1-forward1	BamHI (GGATCC)

Two of the four available PCR primers (Table 5) are used for cloning  
the *at-cbf1* gene into the expression vectors containing each inducible  
promoter described above. The four primers have these sequences: *cbf1*-  
reverse 1 5'-GGAAGATCTTGAAACAGAGTACTCTGATCAATGAACTC-3'  
[SEQ. ID. No. 34], *cbf1*-forward 1 5'-  
CGCGGATCCCTCGTTTCTACAACAATAAAATAAAATG-3' [SEQ. ID.  
No. 35], *cbf1*-reverse 2 5'-

GGGGTACCTGAAACAGAGTACTCTGATCAATGAACTC-3' [SEQ. ID. No. 36],  
and *cbf1*-forward 2 5'-

GCTCTAGACTCGTTTCTACAACAATAAAATAAAATG-3' [SEQ. ID. No.  
37]. For example, for the *Dreb2a*, *P5CS*, and *COR47* promoters that are  
5 ligated to a *Bam*HI and *Bgl*II flanked insert, the *cbf1*-reverse 1 and *cbf1*-forward  
1 primers [SEQ. ID. No. 34 and 35, respectively] are used to isolate *cbf1* gene  
from *Arabidopsis thaliana* genomic DNA. The *cbf1*-reverse primer includes a  
*Bgl*II (AGATCT) restriction site near the 5'-end of the primer and *cbf1*-forward  
primer has a *Bam*HI (GGATCC) restriction site near the 5'-end of the primer. A  
10 PCR product of 764 bp is expected. The genomic DNA (10 ng) is used as a  
template in a PCR reaction under conditions suggested by the manufacturer  
(Boehringer Mannheim). The reaction conditions to be used in this PCR  
experiment are as follows: Segment 1: 94°C, 2 minutes; Segment 2: 94°C, 30  
seconds; 55°C, 1 minute; 72°C, 1 minute, for a total of 35 cycles; Segment 3:  
15 72°C for 10 minutes.

The PCR products are subject to electrophoresis in a 0.8% agarose gel  
and visualized by ethidium bromide staining. The DNA fragment containing  
*cbf1* is excised and purified by using a Qiaquick gel extraction kit (Qiagen, CA).  
The purified fragment and the vector pMBI2001 containing the inducible  
20 promoter (Table 5) are each digested with *Bgl*II and *Bam*HI restriction enzymes  
at 37°C for 2 hours. Both DNA samples are purified by using the Qiaquick  
purification kit (Qiagen, CA) and ligated at a ratio of 3:1 (vector to insert).  
Ligation reactions using T4 DNA ligase (New England Biolabs, MA) are carried  
out at 16°C for 16 hours. The ligated DNAs are transformed into competent  
25 cells of the *E. coli* strain DH5 by using the heat shock method. The  
transformation are plated on LB plates containing 100 (g/ml spectinomycin  
(Sigma).

Individual colonies are grown overnight in five milliliters of LB broth  
containing 100 g/ml spectinomycin at 37°C. Plasmid DNA are purified by using  
30 Qiaquick Mini Prep kits (Qiagen, CA). The presence of the *cbf1* insert is verified  
by restriction mapping with *Bgl*II and *Bam*HI. The plasmid DNA is also subject  
to double-strand DNA sequencing analysis by using vector primer E9.1(5'-  
CAAAGTCAGTAGGATTCTGGTGTGT-3') [SEQ. ID. No. 33]. The other  
primers shown in Table 5 and appropriate restriction enzymes are used in a

similar way to clone the *Cbf1* gene into plasmids containing the other inducible promoters. The resulting plasmids are listed in Table 6 and shown in Figures 17A-17G.

5 A similar cloning strategy may be used to clone other genes, such as *cbf2*, *cbf3*, and the other full length CBF genes listed in Table 9 and shown in figure 18 (new CBF gene table) into plasmids containing inducible promoters.

**Table 6.**

10	<b>Construct name</b>	<b>Promoter name</b>	<b>Figure name</b>
	PMBI2008	Dreb2a	FIGURE 17A
	PMBI2009	P5CS	FIGURE 17B
	PMBI2010	Rd22	FIGURE 17C
	PMBI2011	Rd29a	FIGURE 17D
15	PMBI2012	Rd29b	FIGURE 17E
	PMBI2013	Rab18	FIGURE 17F
	PMBI2014	Cor47	FIGURE 17G

20 (3) *Cloning of the cbf1 gene into the plasmids containing seed-specific promoters*

Several different CBF coding regions with different translational efficiencies in *arabidopsis thaliana* are cloned into expression vectors pMEN1001, pMEN1002, and pMEN1003 to produce different levels of CBF protein in transgenic plants. The 5' and 3' PCR primers used to isolate *cbf1* gene from *arabidopsis thaliana* genomic DNA are listed below.

5'-Primer *cbf5pri.atg.seq* for isolating *cbf1.1* gene [SEQ. ID. No. 96]

5'-ggaagatctatGAAACAGAGTACTCTGATCAATGAACTC-3'

5'-Primer *cbf5pri.wt.seq* for isolating *cbf1.2* gene [SEQ. ID. No. 97]

5'-ggaagatctGAAACAGAGTACTCTGATCAATGAACTC-3'

5'-Primer *cbf5pri.inframeatg.seq* for isolating *cbf1.3* gene [SEQ. ID. No. 98]

5'-ggaagatctatGAACAGAGTACTCTGATCAATGAACTC-3'

5'-Primer cbf5pri.dbatg.seq for isolating *cbf1.4* gene [SEQ. ID. No. 99]  
 5'-ggaagatctatGAACAGAGTACTCTGATgCAATGAACTC-3'

5 3'-Primer cbf1.long3pri.seq for isolating *cbf1.1-4* genes [SEQ. ID. No. 100]  
 5'-ggaggatcCTCGTTTCTACAACAATAAAATAAAATAAAATGAAGGAACC-  
 3'

The *cbf1* gene is cloned into pMEN050 at restriction sites HindIII and  
 10 BamH1 by using a similar strategy as described in subsection (2) of this section  
 for cloning of the *cbf1* gene into the plasmids containing the inducible  
 promoters. The resulting constructs containing *cbf1.1-4* genes are  
 pMEN1001.1-4 plasmids, pMEN1002.1-4 plasmids and pMEN1003.1-4  
 plasmids, respectively. The presence of the *cbf1* gene inserts is verified by  
 15 restriction mapping with HindIII and BamH1 restriction enzymes to cut out the  
 cloned insert. The plasmid DNA is also subject to double-strand DNA  
 sequencing analysis using a vector primer (E9.1 primer 5'-  
 CAAACTCAGTAGGATTCTGGTGTGT-3' [SEQ. ID. No. 33].

#### 20 D. Transformation of *Agrobacterium* with Plasmids Containing CBF1 Gene and Inducible or Tissue-Specific Promoters

After the plasmid vectors containing *cbf1* gene and inducible promoters  
 are constructed, these vectors are used to transform *Agrobacterium*  
 25 *tumefaciens* cells expressing the gene products. The stock of *Agrobacterium*  
*tumefaciens* cells for transformation are made as described by Nagel et al.  
 FEMS Microbiol Letts 67: 325-328 (1990). *Agrobacterium* strain ABI is grown  
 in 250 ml LB medium (Sigma) overnight at 28°C with shaking until an  
 absorbance ( $A_{600}$ ) of 0.5 – 1.0 is reached. Cells are harvested by centrifugation  
 30 at 4,000 x g for 15 min at 4 °C. Cells are then resuspended in 250 µl chilled  
 buffer (1 mM HEPES, pH adjusted to 7.0 with KOH). Cells are centrifuged  
 again as described above and resuspended in 125 µl chilled buffer. Cells are  
 then centrifuged and resuspended two more times in the same HEPES buffer  
 as described above at a volume of 100 µl and 750 µl, respectively.

Resuspended cells are then distributed into 40 µl aliquots, quickly frozen in liquid nitrogen, and stored at -80°C.

*Agrobacterium* cells are transformed with plasmids formed as described above in Section 4B(2) following the protocol described by Nagel et al. FEMS Microbiol Letts **67**: 325-328 (1990). For each DNA construct to be transformed, 50 – 100 ng DNA (generally resuspended in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0) is mixed with 40 µl of *Agrobacterium* cells. The DNA/cell mixture is then transferred to a chilled cuvette with a 2mm electrode gap and subject to a 2.5 kV charge dissipated at 25 µF and 200 µF using a Gene Pulser II apparatus (Bio-Rad). After electroporation, cells are immediately resuspended in 1.0 ml LB and allowed to recover without antibiotic selection for 2 – 4 hours at 28 °C in a shaking incubator. After recovery, cells are plated onto selective medium of LB broth containing 100 µg/ml spectinomycin (Sigma) and incubated for 24-48 h at 28 °C. Single colonies are then picked and inoculated in fresh medium. The presence of the plasmid construct are verified by PCR amplification and sequence analysis.

#### **E. Transformation of *Arabidopsis* Plants with *Agrobacterium tumefaciens* Carrying Expression Vector for CBF1 Protein**

After transformation of *Agrobacterium tumefaciens* with plasmid vectors containing *cbf1* gene and inducible promoters, single *Agrobacterium* colonies containing each of pMBI2008 - pMBI2014 are identified, propagated, and used to transform *Arabidopsis* Plants. Briefly, 500 ml cultures of LB medium containing 100ug/ml spectinomycin are inoculated with the colonies and grown at 28 C with shaking for 2 days until an absorbance ( $A_{600}$ ) of > 2.0 is reached. Cells are then harvested by centrifugation at 4,000 x g for 10 min, and resuspended in infiltration medium (1/2 X Murashige and Skoog salts (Sigma), 1 X Gamborg's B-5 vitamins (Sigma), 5.0% (w/v) sucrose (Sigma), 0.044 µM benzylamino purine (Sigma), 200 µl/L Silwet L-77 (Lehle Seeds) until an absorbance ( $A_{600}$ ) of 0.8 is reached.

Prior to transformation, *Arabidopsis thaliana* seeds (ecotype Columbia) are sown at a density of ~10 plants per 4" pot onto Pro-Mix BX potting medium (Hummert International) covered with fiberglass mesh (18 mm X 16 mm).

Plants are grown under continuous illumination (50-75 µE/m<sup>2</sup>/sec) at 22-23 C

with 65-70% relative humidity. After about 4 weeks, primary inflorescence stems (bolts) are cut off to encourage growth of multiple secondary bolts. After flowering of the mature secondary bolts, plants are prepared for transformation by removal of all siliques and opened flowers.

- 5           The pots are then immersed upside down in the mixture of Agrobacterium / infiltration medium as described above for 30 sec, and placed on their sides to allow draining into a 1' x 2' flat surface covered with plastic wrap. After 24 h, the plastic wrap is removed and pots are turned upright. The immersion procedure is repeated one week later, for a total of two immersions
- 10          per pot. Seeds are then collected from each transformation pot and analyzed following the protocol described below.

#### F. Identification of Arabidopsis Primary Transformants

- 15           Seeds collected from the transformation pots are sterilized essentially as follows. Seeds are dispersed into in a solution containing 0.1% (v/v) Triton X-100 (Sigma) and sterile H<sub>2</sub>O and washed by shaking the suspension for 20 min. The wash solution is then drained and replaced with fresh wash solution to wash the seeds for 20 min with shaking. After removal of the second wash
- 20          solution, a solution containing 0.1% (v/v) Triton X-100 and 70% EtOH (Equistar) is added to the seeds and the suspension is shaken for 5 min. After removal of the ethanol/detergent solution, a solution containing 0.1% (v/v) Triton X-100 and 30% (v/v) bleach (Chlorox) is added to the seeds, and the suspension is shaken for 10 min. After removal of the bleach/detergent
- 25          solution, seeds are then washed five times in sterile distilled H<sub>2</sub>O. The seeds are stored in the last wash water at 4 °C for 2 days in the dark before being plated onto antibiotic selection medium (1 X Murashige and Skoog salts (pH adjusted to 5.7 with 1M KOH), 1 X Gamborg's B-5 vitamins, 0.9% phytagar (Life Technologies), and 50 µg/L kanamycin). Seeds are germinated under
- 30          continuous illumination (50-75 µE/m<sup>2</sup>/sec) at 22-23 °C. After 7-10 days of growth under these conditions, kanamycin resistant primary transformants (T<sub>1</sub> generation) are visible and are obtained for each of constructs pMBI2008 - pMBI2014. These seedlings are transferred first to fresh selection plates where the seedlings continued to grow for 3-5 more days, and then to soil (Pro-

Mix BX potting medium). Progeny seeds ( $T_2$ ) are collected; kanamycin resistant seedlings selected and analyzed as described above.

#### G. Transformation of Cereal Plants with Plasmid Vectors Containing *cbf1* Gene and Inducible Promoters

Cereal plants, such as corn, wheat, rice, sorghum and barley, can also be transformed with the plasmid vectors containing the *cbf* genes and inducible promoters to increase their tolerance to environmental stresses. In these cases, the cloning vector, pMEN020, is modified to replace the NptII coding region with the BAR gene of *Streptomyces hygroscopicus* that confers resistance to phosphinothricin. The KpnI and BglII sites of the Bar gene are removed by site-directed mutagenesis with silent codon changes. After cloning of the inducible promoters into the modified plasmid by the same procedures described above, the *at-cbf* coding region of *cbf1* gene is inserted into the plasmid following the same procedures as described above. The resulted plasmids are listed in Table 7.

Table 7

Promoter name	Construct name
Dreb2a	PMBI2015
P5CS	PMBI2016
Rd22	PMBI2017
Rd29a	PMBI2018
Rd29b	PMBI2019
Rab18	PMBI2020
Cor47	PMBI2021

It is now routine to produce transgenic plants of most cereal crops (Vasil, I., Plant Molec. Biol. **25**: 925-937 (1994)) such as corn, wheat, rice, sorghum (Cassas, A. et al., Proc. Natl. Acad Sci USA **90**: 11212-11216 (1993) and barley (Wan, Y. and Lemeaux, P. Plant Physiol. **104**:37-48 (1994) Other direct DNA transfer methods such as the microprojectile gun or *Agrobacterium tumefaciens*-mediated transformation can be used for corn (Fromm. et al. Bio/Technology **8**: 833-839 (1990); Gordon-Kamm et al. Plant Cell **2**: 603-618

(1990); Ishida, Y., *Nature Biotechnology* **14**:745-750 (1990)), wheat (Vasil, et al. *Bio/Technology* **10**:667-674 (1992) ; Vasil et al., *Bio/Technology* **11**:1553-1558 (1993); Weeks et al., *Plant Physiol.* **102**:1077-1084 (1993)), rice (Christou *Bio/Technology* **9**:957-962 (1991); Hiei et al. *Plant J.* **6**:271-282 (1994); Aldemita and Hodges, *Planta* **199**:612-617; Hiei et al., *Plant Mol Biol.* **35**:205-18 (1997)). For most cereal plants, embryogenic cells derived from immature scutellum tissues are the preferred cellular targets for transformation (Hiei et al., *Plant Mol Biol.* **35**:205-18 (1997); Vasil, *Plant Molec. Biol.* **25**: 925-937 (1994)).

Plasmids according to the present invention may be transformed into corn embryogenic cells derived from immature scutellar tissue by using microprojectile bombardment, with the A188XB73 genotype as the preferred genotype (Fromm, et al., *Bio/Technology* **8**: 833-839 (1990); Gordon-Kamm et al., *Plant Cell* **2**: 603-618 (1990)). After microprojectile bombardment the tissues are selected on phosphinothricin to identify the transgenic embryogenic cells (Gordon-Kamm et al., *Plant Cell* **2**: 603-618 (1990)). Transgenic plants are regenerated by standard corn regeneration techniques (Fromm, et al., *Bio/Technology* **8**: 833-839 (1990); Gordon-Kamm et al., *Plant Cell* **2**: 603-618 (1990)).

The plasmids prepared as described above can also be used to produce transgenic wheat and rice plants (Christou, *Bio/Technology* **9**:957-962 (1991); Hiei et al., *Plant J.* **6**:271-282 (1994) ; Aldemita and Hodges, *Planta* **199**:612-617 (1996); Hiei et al., *Plant Mol Biol.* **35**:205-18 (1997)) by following standard transformation protocols known to those skilled in the art for rice and wheat (Vasil, et al. *Bio/Technology* **10**:667-674 (1992) ; Vasil et al., *Bio/Technology* **11**:1553-1558 (1993); Weeks et al., *Plant Physiol.* **102**:1077-1084 (1993)), where the BAR gene is used as the selectable marker.

#### **H. Transformation of Plants with Plasmid Vectors Containing *cbf1* Gene and Seed-Specific Promoters**

The binary constructs containing seed-specific napin promoters (pMEN1001.1-4; pMEN1002.1-4; and pMEN1003.1-4) are used to transform canola and rapeseed plants as described (Moloney et al. US Patent No. 5,750,871), except that the Bar gene selectable marker is used. These



constructs are also used to transform regenerable barley cells by microprojectile bombardment (Wan and Lemaux, Plant Physiol. **104**: 37-48 (1994)). After bombardment the tissues are selected on phosphinothricin to identify the transgenic embryogenic cells. Transgenic cells are regenerated by standard barley regeneration techniques (Wan and Lemaux Plant Physiol. **104**: 37-48 (1994)).

## 5. Identification of CBF1 Homologs CBF2 and CBF3 Using CBF1

This example describes two homologs of CBF1 from *Arabidopsis thaliana* and named them CBF2 and CBF3.

*CBF2* and *CBF3* have been cloned and sequenced as described below. The sequences of the DNA and encoded proteins are set forth in SEQ ID NOS: 12, 13, 14 and 15. Figure 12 shows the DNA sequence for *CBF2* encoding CBF2. Figure 13 shows the DNA sequence for *CBF3* encoding CBF3.

A lambda cDNA library prepared from RNA isolated from *Arabidopsis thaliana* ecotype Columbia (Lin and Thomashow, Plant Physiol. **99**: 519-525 (1992)) was screened for recombinant clones that carried inserts related to the *CBF1* gene (Stockinger, E. J., et al., Proc Natl Acad Sci USA **94**:1035-1040 (1997)). *CBF1* was <sup>32</sup>P-radiolabeled by random priming (Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed. 2, Cold Spring Harbor Laboratory Press, New York (1989)) and used to screen the library by the plaque-lift technique using standard stringent hybridization and wash conditions (Hajela, R. K., et al., Plant Physiol **93**:1246-1252 (1990); Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor laboratory Press, New York (1989) 6 X SSPE buffer, 60 °C for hybridization and 0.1 X SSPE buffer and 60 °C for washes). Twelve positively hybridizing clones were obtained and the DNA sequences of the cDNA inserts were determined at the MSU-DOE Plant Research Laboratory sequencing facility. The results indicated that the clones fell into three classes. One class carried inserts corresponding to *CBF1*. The two other classes carried sequences corresponding to two different homologs of *CBF1*, designated *CBF2* and *CBF3*. The nucleic acid sequences and predicted protein coding sequences for *CBF1*, *CBF2* and *CBF3* appear at Figure14.

A comparison of the nucleic acid sequences of *CBF1*, *CBF2* and *CBF3* indicate that they are 83 to 85% identical as shown in Table 8. Figure 14 shows the amino acid alignment of proteins CBF1, CBF2 and CBF3.

5

TABLE 8

	Percent identity <sup>a</sup>	
	DNA <sup>b</sup>	Polypeptide
cbf1/cbf2	85	86
cbf1/cbf3	83	84
cbf2/cbf3	84	85

<sup>a</sup> Percent identity was determined using the *Clustal* algorithm from the Megalign program (DNASTAR, Inc.).

10

<sup>b</sup> Comparisons of the nucleic acid sequences of the open reading frames are shown.

15

Similarly, the amino acid sequences of the three CBF polypeptides range from 84 to 86% identity. An alignment of the three amino acid sequences reveals that most of the differences in amino acid sequence occur in the acidic C-terminal half of the polypeptide. This region of CBF1 serves as an activation domain in both yeast and *Arabidopsis* (not shown).

20

Residues 47 to 106 of CBF1 correspond to the AP2 domain of the protein, a DNA binding motif that to date, has only been found in plant proteins. A comparison of the AP2 domains of CBF1, CBF2 and CBF3 indicates that there are a few differences in amino acid sequence. These differences in amino acid sequence might have an effect on DNA binding specificity.

## 6. Activation of Transcription In Yeast Containing C-repeat/DRE Using CBF1, CBF2 and CBF3

This example shows that CBF1, CBF2 and CBF3 activate transcription in yeast containing CRT/DREs upstream of a reporter gene. The CBFs were expressed in yeast under control of the ADC1 promoter on a 2 $\mu$  plasmid (pDB20.1; Berger, S. L., et al., Cell 70:251-265 (1992)). Constructs expressing the different CBFs were transformed into yeast reporter strains which had the indicated CRT/DRE upstream of the lacZ reporter gene. Copy number of the CRT/DREs and its orientation relative to the direction of transcription from each promoter is indicated by the direction of the arrow.

Figure 15 is a graph showing transcription regulation of CRT/DRE containing reporter genes by CBF1, CBF2 and CBF3 genes in yeast. In Figure 15, the vertical lines across the arrows of the COR15a construct represent the m3cor15a mutant CRT/DRE construct. Each CRT/DRE-lacZ construct was integrated into the URA3 locus of yeast. Error bars represent the standard deviation derived from three replicate transformation events with the same CBF activator construct into the respective reporter strain. Quantitative B-gal assays were performed as described by Rose and Botstein (Rose, M., et al., Methods Enzymol. 101:167-180 (1983)).

## 7. Homologous CBF Encoding Genes In Other Plants.

This example shows that homologous sequences to CBF1 are present in other plants. The presence of these homologous sequences suggest that the same or similar cold regulated environmental stress response regulatory elements such as the C-repeat/DRE of Arabidopsis (CCGAC) exist in other plants. This example serves to indicate that genes with significant homology to CBF1, CBF2 and CBF3 exist in a wide range of plant species.

Total plant DNAs from *Arabidopsis thaliana*, *Nicotiana tabacum*, *Lycopersicon pimpinellifolium*, *Prunus avium*, *Prunus cerasus*, *Cucumis sativus*, and *Oryza sativa* were isolated according to Stockinger et al (Stockinger, E. J., et al., J. Heredity, 87:214-218 (1996)). Approximately 2 to 10  $\mu$ g of each DNA sample was restriction digested, transferred to nylon membrane (Micron Separations, Westboro, MA) and hybridized according to Walling et al.

(Walling, L. L., et al., Nucleic Acids Res. **16**:10477-10492 (1988)).

Hybridization conditions were: 42 °C in 50% formamide, 5X SSC, 20 mM phosphate buffer 1X Denhardt's, 10% dextran sulfate, and 100µg/ml herring sperm DNA. Four low stringency washes at RT in 2X SSC, 0.05% Na sarcosyl and 0.02% Na<sub>4</sub> pyrophosphate were performed prior to high stringency washes at 55 °C in 0.2X SSC, 0.05% Na sarcosyl and 0.01% Na<sub>4</sub> pyrophosphate. High stringency washes were performed until no counts were detected in the washout. The BclI-BglII fragment of CBF1 (Stockinger et al., Proc Natl Acad Sci USA **94**:1035-1040 (1997)) was gel isolated (Sambrook et al., Molecular Cloning. A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, New York (1989)) and direct prime labelled (Feinberg and Vogelstein, Anal. Biochem **132**: 6-13 (1982)) using the primer MT117 (TTGGCGGCTACGAATCCC; SEQ ID NO:16). Specific activity of the radiolabelled fragment was approximately 4 x 10<sup>8</sup> cpm/µg. Autoradiography was performed using HYPERFILM-MP (Amersham) at -80 °C with one intensifying screen for 15 hours.

Autoradiography of the gel showed that DNA sequences from *Arabidopsis thaliana*, *Nicotiana tabacum*, *Lycopersicon pimpinellifolium*, *Prunus avium*, *Prunus cerasus*, *Cucumis sativus*, and *Oryza sativa* hybridized to the labeled BclI, BglII fragment of CBF1. These results suggest that homologous CBF encoding genes are present in a variety of other plants.

## 8. Identification Of Homologous Sequence To CBF1 In Canola

This example describes the identification of homologous sequences to CBF1 in canola using PCR. Degenerate primers were designed for regions of AP2 binding domain and outside of the AP2 (carboxyl terminal domain). More specifically, the following degenerate PCR primers were used:

Mol 368 (reverse) 5'- CAY CCN ATH TAY MGN GGN GT -3'

Mol 378 (forward) 5'- GGN ARN ARC ATN CCY TCN GCC -3'

(Y: C/T, N: A/C/G/T, H: A/C/T, M: A/C, R: A/G )

Primer Mol 368 is in the AP2 binding domain of CBF1 (amino acid seq: H P I Y R G V) while primer Mol 378 is outside the AP2 domain (carboxyl terminal domain)(amino acid seq: M A E G M L L P).

5 The genomic DNA isolated from *Brassica Napus* was PCR amplified by using these primers following these conditions: an initial denaturation step of 2 min at 93 °C; 35 cycles of 93 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min ; and a final incubation of 7 min at 72 °C at the end of cycling.

10 The PCR products were separated by electrophoresis on a 1.2% agarose gel and, transferred to nylon membrane and hybridized with the AT CBF1 probe prepared from Arabidopsis genomic DNA by PCR amplification. The hybridized products were visualized by colormetric detection system (Boehringer Mannheim) and the corresponding bands from a similar agarose gel were isolated (By Qiagen Extraction Kit). The DNA fragments were ligated into the TA clone vector from TOPO TA Cloning Kit (Invitrogen) and  
15 transformed into E. coli strain TOP10 (Invitrogen).

Seven colonies were picked and the inserts were sequenced on an ABI 377 machine from both strands of sense and antisense after plasmid DNA isolation. The DNA sequence was edited by sequencer and aligned with the AtCBF1 by GCG software and NCBI blast searching.

20 Figure 16 shows an amino acid sequence of a homolog [CAN1; SEQ. ID. No. 17] identified by this process and its alignment to the amino acid sequence of CBF1. The nucleic acid sequence for CAN1 is listed herein as SEQ. ID. No. 18.

As illustrated in Figure 16, the DNA sequence alignment in four regions  
25 of BN-CBF1 shows 82% identity in the AP2 binding domain region and range from 75% to 83% with some alignment gaps due to regions of lesser homology or introns in the genomic sequence. The aligned amino acid sequences show that the BNCBF1 gene has 88% identity in the AP2 domain region and 85% identity outside the AP2 domain when aligned for two insertion sequences that  
30 are outside the AP2 domain. The extra amino acids in the 2 insertion regions are either due to the presence of introns in this region of the BNCBF1 gene, as it was derived from genomic DNA, or could be due to extra amino acids in these regions of the BNCBF1 gene. Isolation and sequencing of a cDNA of the BNCBF1 gene using the genomic DNA as a probe will resolve this.

### 9. Identification Of Homologous Sequences To CBF1 In Canola and other Species

A PCR strategy similar to that described in Example 8 was used to isolate additional CBF homologues from *Brassica juncea*, *Brassica napus*, *Brassica oleracea*, *Brassica rapa*, *Glycine max*, *Raphanus sativus* and *Zea Maize*. The nucleotide (e.g. bjCBF1) and peptide sequences (e.g. BJCBF1-PEP) of these isolated CBF homologues are shown in Figures 18A and 18B, respectively. Table 9 lists the sequence names and sequence ID Nos. of these isolated CBF homologues. The PCR primers are internal to the gene so partial gene sequences are initially obtained. The full length sequences of some of these genes were further isolated by inverse PCR or ligated linker PCR. One skilled in the art can use the conserved regions in these genes to design PCR primers to isolate additional CBF genes.

**Table 9**

DNA Seq. Name	Seq. ID No.	Peptide Seq. Name	Seq. ID No.
bjCBF1	38	BJCBF1-PEP	39
bjCBF2	40	BJCBF2-PEP	41
bjCBF3	42	BJCBF3-PEP	43
bjCBF4	44	BJCBF4-PEP	45
bnCBF1	46	BNCBF1-PEP	47
bnCBF2	48	BNCBF2-PEP	49
bnCBF3	50	BNCBF3-PEP	51
bnCBF4	52	BNCBF4-PEP	53
bnCBF5	54	BNCBF5-PEP	55
bnCBF6	56	BNCBF6-PEP	57
bnCBF7	58	BNCBF7-PEP	59
bnCBF8	60	BNCBF8-PEP	61
bnCBF9	62	BNCBF9-PEP	63
boCBF1	64	BOCBF1-PEP	65
boCBF2	66	BOCBF2-PEP	67
boCBF3	68	BOCBF3-PEP	69

DNA Seq. Name	Seq. ID No.	Peptide Seq. Name	Seq. ID No.
boCBF4	70	BOCBF4-PEP	71
boCBF5	72	BOCBF5-PEP	73
brCBF1	74	BRCBF1-PEP	75
brCBF2	76	BRCBF2-PEP	77
brCBF3	78	BRCBF3-PEP	79
brCBF4	80	BRCBF4-PEP	81
brCBF5	82	BRCBF5-PEP	83
brCBF6	84	BRCBF6-PEP	85
brCBF7	86	BRCBF7-PEP	87
gmCBF1	88	GMCBF1-PEP	89
rsCBF1	90	RSCBF1-PEP	91
rsCBF2	92	RSCBF2-PEP	93
zmCBF1	94	ZMCBF1-PEP	95

Figure 19A shows an amino acid alignment of the AP2 domains of the CBF proteins listed in Table 9 with their consensus sequences highlighted. Figure 19A also provides a comparison of the consensus sequence with that of the tobacco DNA binding protein EREBP2 (Okme-Takagi, M., et al., The Plant Cell 7:173-182 (1995). The sequences of these CBF proteins are BRCBF3-PEP [SEQ. ID. No. 79], BRCBF6-PEP [SEQ. ID. No. 85], BNCBF5-PEP [SEQ. ID. No. 55], ATCBF2-PEP [SEQ. ID. No. 13], ATCBF3-PEP [SEQ. ID. No. 15], ATCBF1-PEP [SEQ. ID. No. 2], BNCBF2-PEP [SEQ. ID. No. 49], BNCBF6-PEP [SEQ. ID. No. 57], BOCBF3-PEP [SEQ. ID. No. 69], BNCBF3-PEP [SEQ. ID. No. 51], BNCBF8-PEP [SEQ. ID. No. 61], BNCBF9-PEP [SEQ. ID. No. 63], BRCBF2-PEP [SEQ. ID. No. 77], BOCBF5-PEP [SEQ. ID. No. 73], BOCBF2-PEP [SEQ. ID. No. 67], RSCBF2-PEP [SEQ. ID. No. 93], BNCBF4-PEP [SEQ. ID. No. 53], BNCBF7-PEP [SEQ. ID. No. 59], BOCBF4-PEP [SEQ. ID. No. 71], BRCBF7-PEP [SEQ. ID. No. 87], BRCBF4-PEP [SEQ. ID. No. 81], BRCBF5-PEP [SEQ. ID. No. 83], RSCBF1-PEP [SEQ. ID. No. 91], BJCBF2-PEP [SEQ. ID. No. 41], BJCBF3-PEP [SEQ. ID. No. 43], BNCBF1-PEP [SEQ. ID. No. 47], BOCBF1-PEP [SEQ. ID. No. 65], BRCBF1-PEP [SEQ. ID. No. 75], BJCBF4-

PEP [SEQ. ID. No. 45], ZMCBF1-PEP [SEQ. ID. No. 95], and GMCBF1-PEP [SEQ. ID. No. 89].

As can be seen from the consensus sequence shown in Figure 19A, a significant portion of the AP2 domain is conserved among the different CBF proteins. In view of this data, Applicants use the conserved sequence in the AP2 domain to define a class of AP2 domain proteins comprising this conserved sequence.

Figure 19B shows an amino acid alignment of the AP2 domains shown in Figure 19A and dreb2a and dreb2b and a consensus sequence between the proteins highlighted. As can be seen, a very high degree of homology exists between AP2 domains shown in Figure 19A and dreb2a and dreb2b. Applicants employ the conserved sequence in the AP2 domain shown in Figure 19B to define a broader class of AP2 domain proteins which are capable of binding to CCG regulatory region.

Figure 19C shows an amino acid alignment of the AP2 domains shown in Figure 19B and tiny and a consensus sequence between the proteins highlighted. As can be seen, a very high degree of homology exists between AP2 domains shown in Figure 19A, dreb2a, dreb2b and tiny. Applicants employ the conserved sequence in the AP2 domain shown in Figure 19C to define a yet broader class of AP2 domain proteins which are capable of binding to CCG regulatory region.

Figure 19D shows a consensus sequence corresponding to the difference between the consensus sequence shown in Figures 19A and tiny. Applicants employ the highlighted portion of the conserved sequence shown in Figure 19D to define a group of amino acid residues which may be critical to binding to a CCG regulatory region.

Figure 19E shows a consensus sequence corresponding to the difference between the consensus sequence shown in Figures 19B and tiny. Applicants employ the highlighted portion of the conserved sequence shown in Figure 19E to define another group of amino acid residues which may be critical to binding to a CCG regulatory region.

Figure 20 shows the amino acid alignment of the amino terminus of the CBF proteins with their consensus sequence highlighted. The sequences of these CBF proteins are: BRCBF3-PEP [SEQ. ID. No. 79], BRCBF6-PEP [SEQ.



ID. No.85], BNCBF5-PEP [SEQ. ID. No. 55], ATCBF2-PEP [SEQ. ID. No. 13],  
 ATCBF3-PEP [SEQ. ID. No. 15], ATCBF1-PEP [SEQ. ID. No. 2], BNCBF2-  
 PEP [SEQ. ID. No. 49], BNCBF6-PEP [SEQ. ID. No. 57], BOCBF3-PEP [SEQ.  
 ID. No. 69], BNCBF3-PEP [SEQ. ID. No. 51], BNCBF8-PEP [SEQ. ID. No. 61],  
 5 BNCBF9-PEP [SEQ. ID. No. 63], BRCBF2-PEP [SEQ. ID. No. 77], BOCBF5-  
 PEP [SEQ. ID. No. 73], BOCBF2-PEP [SEQ. ID. No. 67], RSCBF2-PEP [SEQ.  
 ID. No. 93], BNCBF4-PEP [SEQ. ID. No. 53 ], BNCBF7-PEP [SEQ. ID. No. 59],  
 BOCBF4-PEP [SEQ. ID. No. 71], BRCBF7-PEP [SEQ. ID. No. 87], BRCBF4-  
 PEP [SEQ. ID. No. 81], BRCBF5-PEP [SEQ. ID. No. 83], and RSCBF1-PEP  
 10 [SEQ. ID. No. 91].

As can be seen from the consensus sequence shown in Figure 20, a  
 significant portion of the amino terminus of CBF proteins is conserved among  
 the different CBF proteins. In view of this data, Applicants employ the  
 conserved sequence in the amino terminus domain to define a class of proteins  
 15 comprising this conserved sequence.

Figure 21A shows the amino acid alignment of the carboxy terminus of  
 24 CBF proteins with their consensus sequences highlighted. The sequences  
 of these CBF proteins are: BRCBF6-PEP [SEQ. ID. No.85], BNCBF5-PEP  
 [SEQ. ID. No. 55], ATCBF2-PEP [SEQ. ID. No. 13], ATCBF3-PEP [SEQ. ID.  
 20 No. 15], ATCBF1-PEP [SEQ. ID. No. 2], BNCBF2-PEP [SEQ. ID. No. 49],  
 BNCBF6-PEP [SEQ. ID. No. 57], BOCBF3-PEP [SEQ. ID. No. 69], BNCBF3-  
 PEP [SEQ. ID. No. 51], BNCBF8-PEP [SEQ. ID. No. 61], BNCBF9-PEP [SEQ.  
 ID. No. 63], BRCBF2-PEP [SEQ. ID. No. 77], BOCBF5-PEP [SEQ. ID. No. 73],  
 RSCBF2-PEP [SEQ. ID. No. 93], BNCBF4-PEP [SEQ. ID. No. 53], BNCBF7-  
 25 PEP [SEQ. ID. No. 59], BOCBF4-PEP [SEQ. ID. No. 71], BRCBF7-PEP [SEQ.  
 ID. No. 87], BRCBF5-PEP [SEQ. ID. No. 83], RSCBF1-PEP [SEQ. ID. No. 91],  
 BJCBF2-PEP [SEQ. ID. No. 41], BJCBF3-PEP [SEQ. ID. No. 43], BNCBF1-  
 PEP [SEQ. ID. No. 47], and BOCBF1-PEP [SEQ. ID. No. 65].

As can be seen from the consensus sequence shown in Figure 21A, a  
 30 significant portion of the carboxy terminus of CBF proteins is conserved among  
 the different CBF proteins. In view of this data, Applicants employ the  
 conserved sequence in the carboxy terminus domain to define a class of  
 proteins comprising this conserved sequence.

Figure 21B shows the amino acid alignment of the carboxy terminus of 9 CBF proteins with their consensus sequences highlighted. The sequences of these CBF proteins are: BNCBF2-PEP [SEQ. ID. No. 49], BOCBF3-PEP [SEQ. ID. No. 69], BNCBF3-PEP [SEQ. ID. No. 51], BNCBF8-PEP [SEQ. ID. No. 61],  
5 BNCBF9-PEP [SEQ. ID. No. 63], BRCBF2-PEP [SEQ. ID. No. 77], BOCBF5-PEP [SEQ. ID. No. 73], BNCBF1-PEP [SEQ. ID. No. 47], and BNCBF6-PEP [SEQ. ID. No. 57].

As can be seen from the consensus sequence shown in Figure 21B, a greater portion of the carboxy terminus is conserved when these 9 CBF  
10 proteins are used. In view of this data, Applicants employ the conserved sequence in the carboxy terminus domain to define another class of proteins comprising this conserved sequence.

While the present invention is disclosed by reference to the preferred embodiments and examples detailed above, it is to be understood that these  
15 examples are intended in an illustrative rather than limiting sense, as it is contemplated that modifications will readily occur to those skilled in the art, which modifications will be within the spirit of the invention and the scope of the appended claims.

## WE CLAIM:

- 1        1.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence sufficiently homologous to the consensus sequence  
3        shown in Figure 19A that the binding protein is capable of binding to a CCG  
4        regulatory sequence.
- 1        2.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence sufficiently homologous the consensus sequence shown  
3        in Figure 19B that the binding protein is capable of binding to a CCG regulatory  
4        sequence.
- 1        3.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence sufficiently homologous the consensus sequence shown  
3        in Figure 19C that the binding protein is capable of binding to a CCG regulatory  
4        sequence.
- 1        4.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence shown in Figure 19A.
- 1        5.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence shown in Figure 19B.
- 1        6.        A binding protein other than CBF-1 in isolated form comprising a  
2        consensus sequence shown in Figure 19C.
- 1        7.        A binding protein other than CBF-1 in isolated form comprising a  
2        sequence selected from the group consisting of an AP2 domain of SEQ. I.D.  
3        Nos. 13, 15, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71,  
4        73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, and 95.
- 1        8.        A non-naturally occurring binding protein comprising an amino acid  
2        sequence capable of binding to a CCG regulatory sequence and an amino acid  
3        sequence which forms a transcription activation region.

- 1 9. DNA in isolated form comprising a sequence encoding any one of the  
2 binding proteins of claims 1-8.
- 1 10. A nucleic acid construct capable of transforming a plant comprising a  
2 sequence encoding any one of the binding proteins of claims 1-8.
- 1 11. A nucleic acid construct according to claim 10, further comprising a  
2 promoter which regulates expression of the binding protein.
- 1 12. A nucleic acid construct according to claim 11, wherein the promoter is  
2 a tissue specific promoter.
- 1 13. A nucleic acid construct according to claim 11, wherein the promoter is  
2 a flower, fruit or seed specific promoter.
- 1 14. A binding protein according to any one of claims 1-8 wherein the  
2 binding protein is a recombinant binding protein expressed within a plant.
- 1 15. A recombinant cell comprising a recombinant binding protein according  
2 to any one of claims 1-8.
- 1 16. A recombinant cell comprising a recombinant sequence encoding a  
2 binding protein according to any one of claims 1-8.
- 1 17. A recombinant cell according to claim 16, further comprising a promoter  
2 which regulates expression of the binding protein.
- 1 18. A recombinant cell according to claim 16, wherein the promoter is a  
2 tissue specific promoter.
- 1 19. A recombinant cell according to claim 16, wherein the promoter is a  
2 flower, fruit or seed specific promoter.
- 1 20. Plant material comprising a recombinant binding protein according to  
2 any one of claims 1-8.

- 1 21. Plant material according to claim 20 wherein the plant material is  
2 selected from the group consisting of seeds, flowers, fruits, leaves and roots of  
3 a plant.
- 1 22. Plant material comprising a recombinant sequence encoding a binding  
2 protein according to any one of claims 1-8.
- 1 23. Plant material according to claim 22, further comprising a promoter  
2 which regulates expression of the binding protein.
- 1 24. Plant material according to claim 22, wherein the promoter is a tissue  
2 specific promoter.
- 1 25. Plant material according to claim 22, wherein the promoter is a flower,  
2 fruit or seed specific promoter.
- 1 26. Plant material according to any one of claims 20-25 wherein the plant  
2 material is a living plant.
- 1 27. A composition of matter derived from plant material according to any  
2 one of claims 20-26.
- 1 28. A method for altering an environmental stress tolerance of a plant  
2 comprising:  
3 transforming a plant with a sequence encoding a binding protein  
4 according to any one of claims 1-8;  
5 expressing the binding protein within the plant; and  
6 stimulating expression of at least one environmental stress tolerance  
7 gene through binding of the binding protein to a DNA regulatory sequence  
8 comprising CCG which regulates expression of one or more environmental  
9 stress tolerance genes in the plant.
- 1 29. A method according to claim 28 wherein the promoter is selected such  
2 that the plant expresses the one or more environmental stress tolerance genes  
3 at a level which differs from a level at which the one or more environmental  
4 stress tolerance genes are expressed in the plant's native state.

- 1        30.     A method according to claim 28 wherein the promoter is selected such
- 2        that the plant expresses the one or more environmental stress tolerance genes
- 3        under environmental conditions at which the plant does not express the binding
- 4        protein in the plant's native state.

1/45

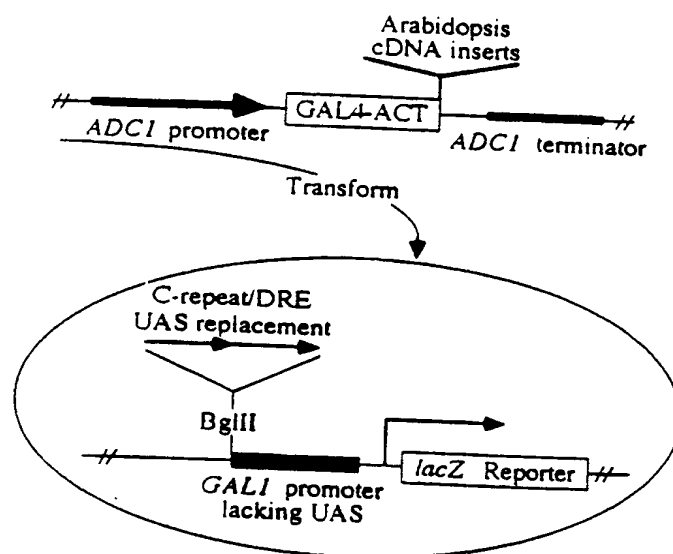


FIGURE 1A

Activity of "positive" plasmids in reporter strains

Oligo	UAS Replacement Sequence		Yeast colony color on X-gal filters
	C-repeat/DRE	Inserts	
MT50	<i>COR15a</i>	→→→→→	Blue
MT50	<i>COR15a</i>	→→→→→	Blue
MT66	<i>COR78</i>	→→→	Blue
MT52	M1 <i>COR15a</i>	→→→	White

FIGURE 1B

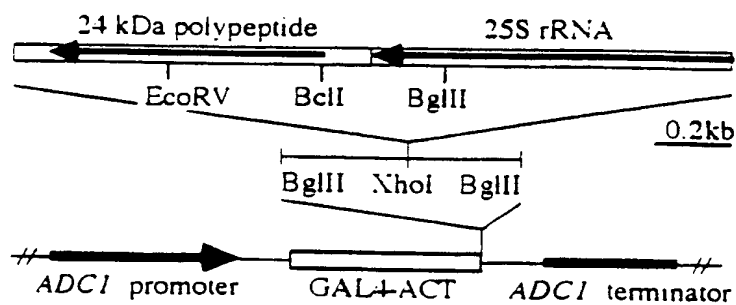


FIGURE 2A



AAAAAGAATCTACCTGAAAAAGAAAAAGAGAGAGAGATATAAATAGCTTACCAAGACAGATATACTATC	71
TTTTATTAAATCCAAAAAGACTGAGAACTCTAGTAACTACGTACTACTTAAACCTTATCCAGTTTCTTGAAA	142
CAGAGTACTCTGATCAATG AAC TCA TTT TCA GCT TTT TCT GAA ATG TTT GGC TCC GAT	200
M N S F S A F S E M F G S D	14
TAC GAG CCT CAA GGC GGA GAT TAT TGT CCG ACG TTG GCC ACG AGT TGT CCG AAG	254
Y E P Q G G D Y C P T L A T S C P K	32
AAA CCG GCG GGC CGT AAG AAG TTT CGT GAG ACT CGT CAC CCA ATT TAC AGA GGA	308
K P A G R K K F R E T R H P I Y R G	50
GTT CGT CAA AGA AAC TCC GGT AAG TGG GTT TCT GAA GTG AGA GAG CCA AAC AAG	362
V R Q R N S G K W V S E V R E P N K	68
AAA ACC AGG ATT TGG CTC GGG ACT TTC CAA ACC GCT GAG ATG GCA GCT CGT GCT	416
K T R I W L G T F Q T A E M A A R A	86
CAC GAC GTC GCT GCA TTA GCC CTC CGT GGC CGA TCA GCA TGT CTC AAC TTC GCT	470
H D V A A L A L R G R S A C L N F A	104
GAC TCG GCT TGG CCG CTA CGA ATC CCG GAG TCA ACA TGC GCC AAG GAT ATC CAA	524
D S A W R L R I P E S T C A K D I Q	122
AAA GCG GCT GCT GAA GCG GCG TTG GCT TTT CAA GAT GAG ACG TGT GAT ACG ACG	578
K A A A E A A L A F Q D E T C D T T	140
ACC ACG GAT CAT GGC CTG GAC ATG GAG GAG ACG ATG GTG GAA GCT ATT TAT ACA	632
T T D H G L D M E E T M V E A I Y T	158
CCG GAA CAG AGC GAA GGT GCG TTT TAT ATG GAT GAG GAG ACA ATG TTT GGG ATG	686
P B Q S E G A F Y M D E E T M F G M	176
CCG ACT TTG TTG GAT AAT ATG GCT GAA GGC ATG CTT TTA CCG CCG CCG TCT GTT	740
P T L L D N M A E G M L L P P P S V	194
CAA TGG AAT CAT AAT TAT GAC GGC GAA GGA GAT GGT GAC GTG TCG CTT TGG AGT	794
Q W N H N Y D G E G D G D V S L W S	212
TAC TAA TATTCGATAGTCGTTTCCATTTTGTACTATAGTTTGAAAATATTCTAGTTCCTTTTGTAGAA	863
Y	213
TGGTTCCTTCATTTTATTTTATTTTATTGTTGTAGAAACGAG	905

FIGURE 2B

SUBSTITUTE SHEET (RULE 26)

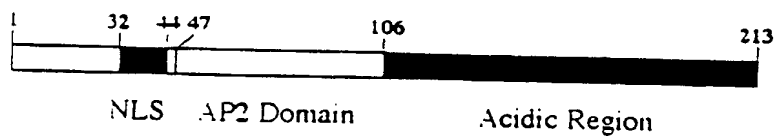


FIGURE 2C

CBF1	IYRGVRQRNSGKQVSEVREPNNKT.RIWLGT	76
EREBP2	HYRGVRQRPWGKFAAEIRDPAKNGARVWLGT	98
CBF1	FOTAEMAARAHQVAALALGRSACLNFADS	106
	: :	
EREBP2	YETAEEAALAYDKAAYRMRGSKALLNFPHR	158

FIGURE 2D

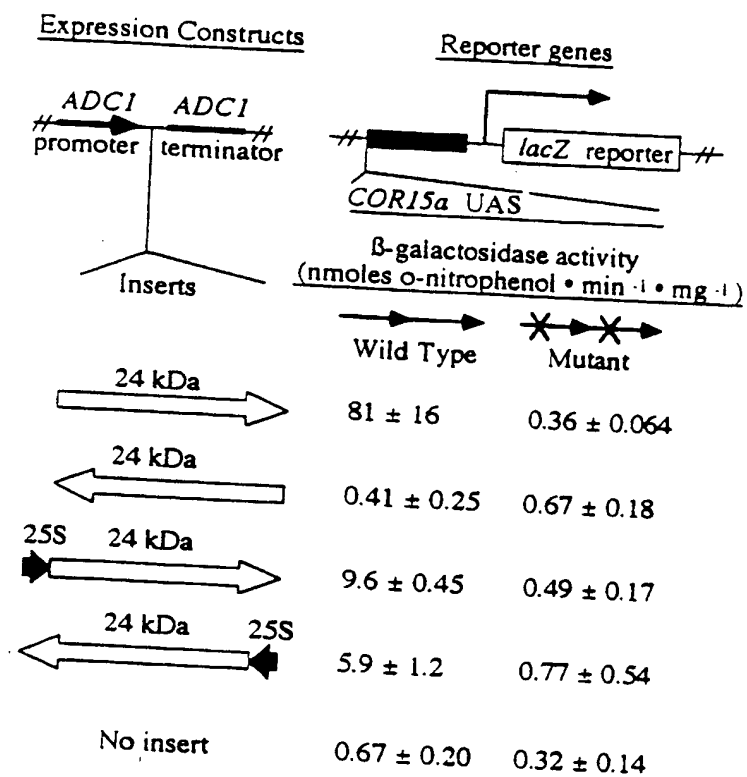


FIGURE 3

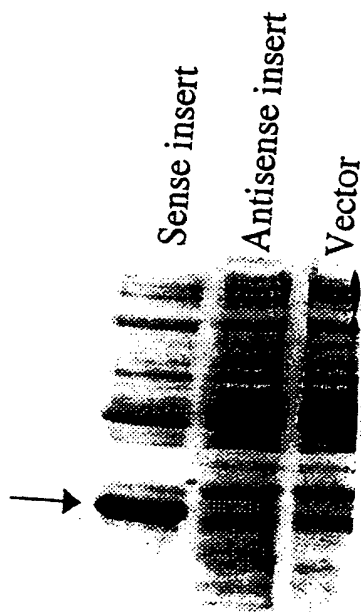


FIGURE 4

SUBSTITUTE SHEET (RULE 26)

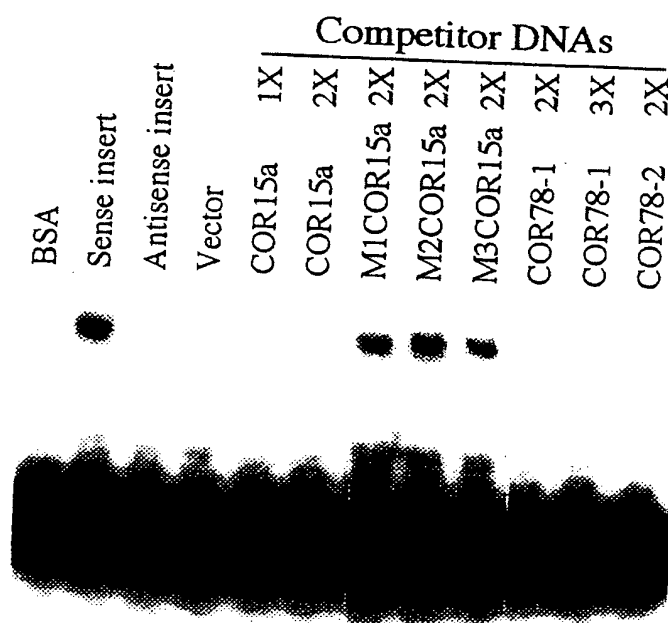


FIGURE 5

**SUBSTITUTE SHEET (RULE 26)**

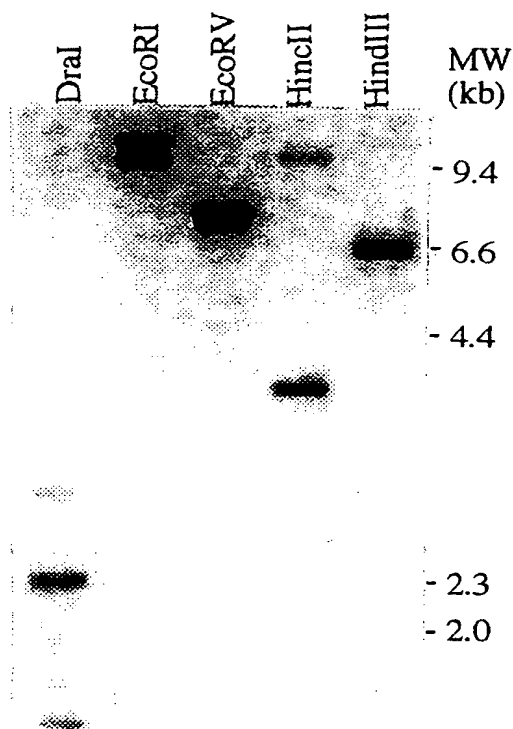


FIGURE 6

SUBSTITUTE SHEET (RULE 26)

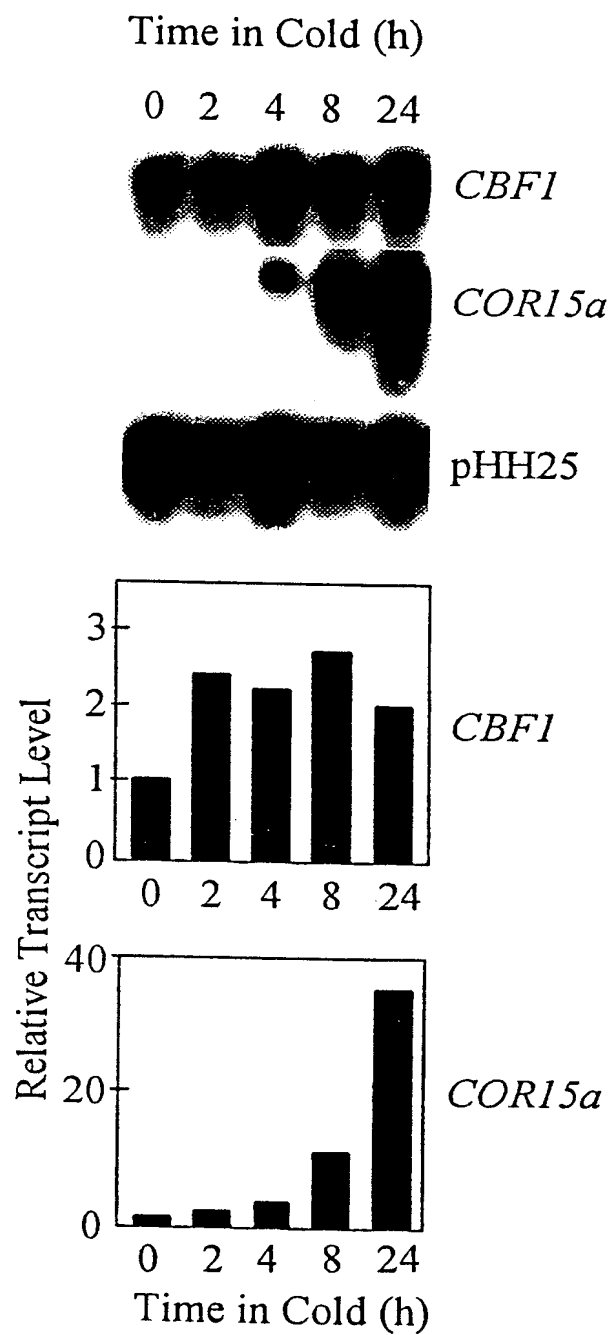


FIGURE 7

SUBSTITUTE SHEET (RULE 26)

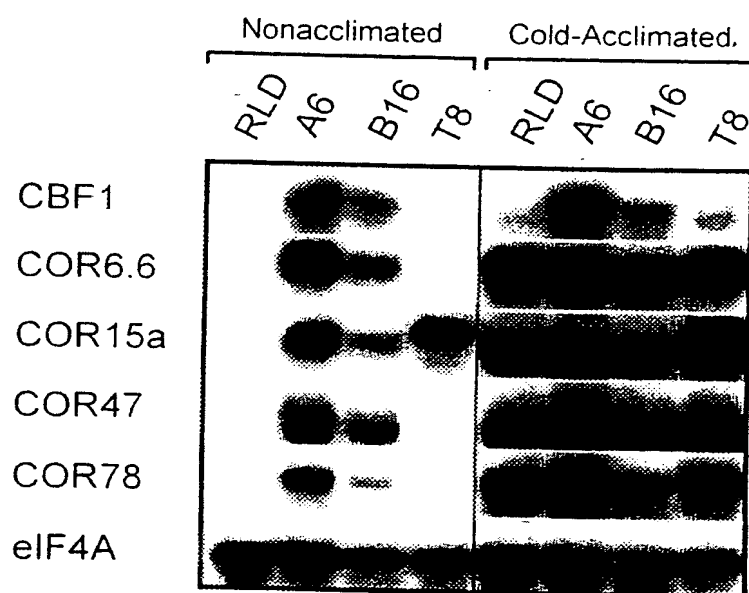


FIGURE 8

SUBSTITUTE SHEET (RULE 26)



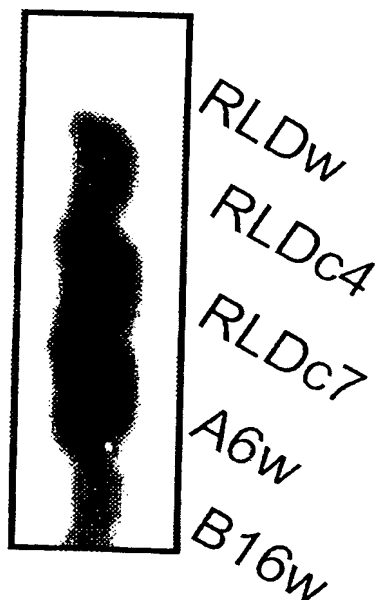


FIGURE 9

SUBSTITUTE SHEET (RULE 26)

12/45

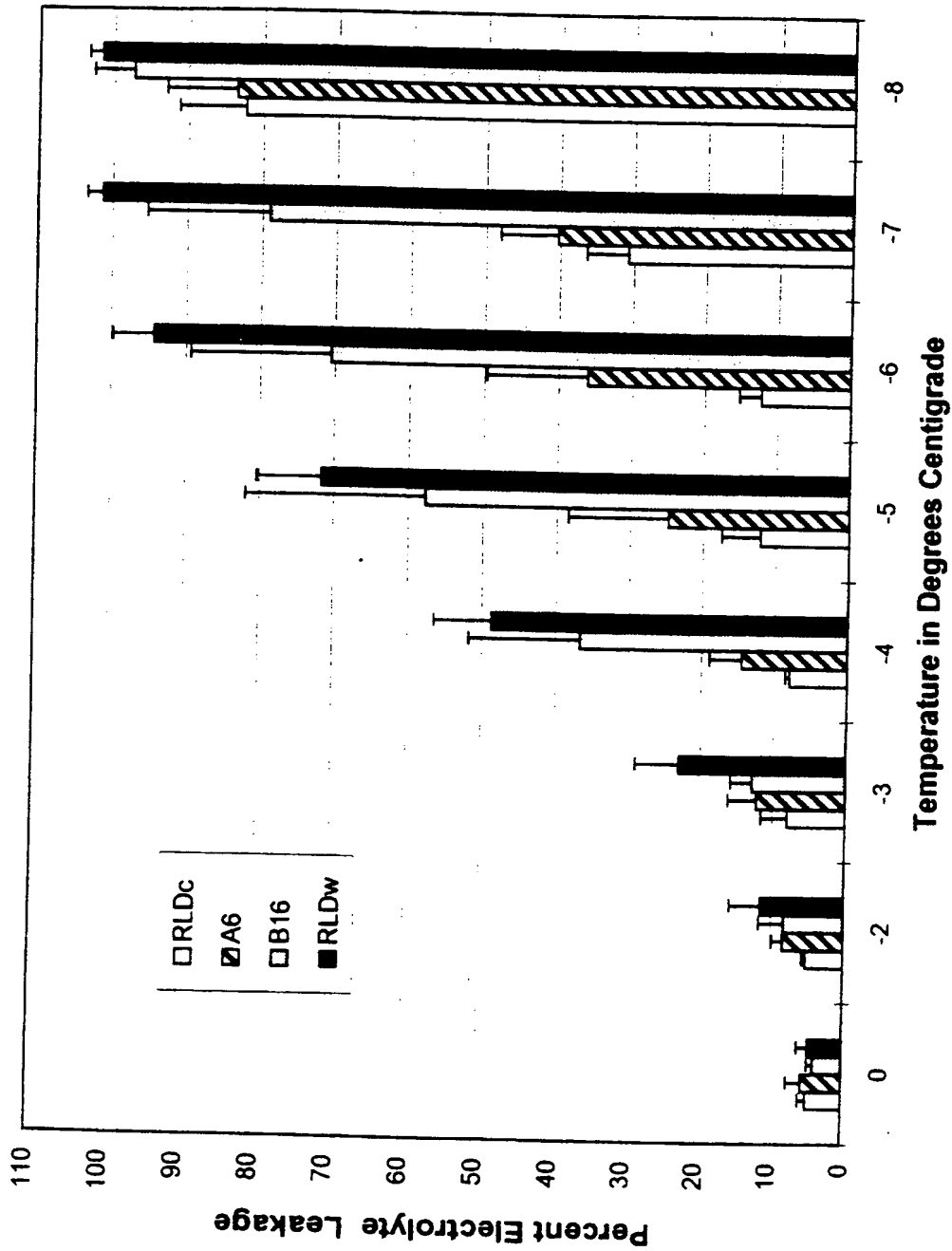


FIGURE 10A

13/45

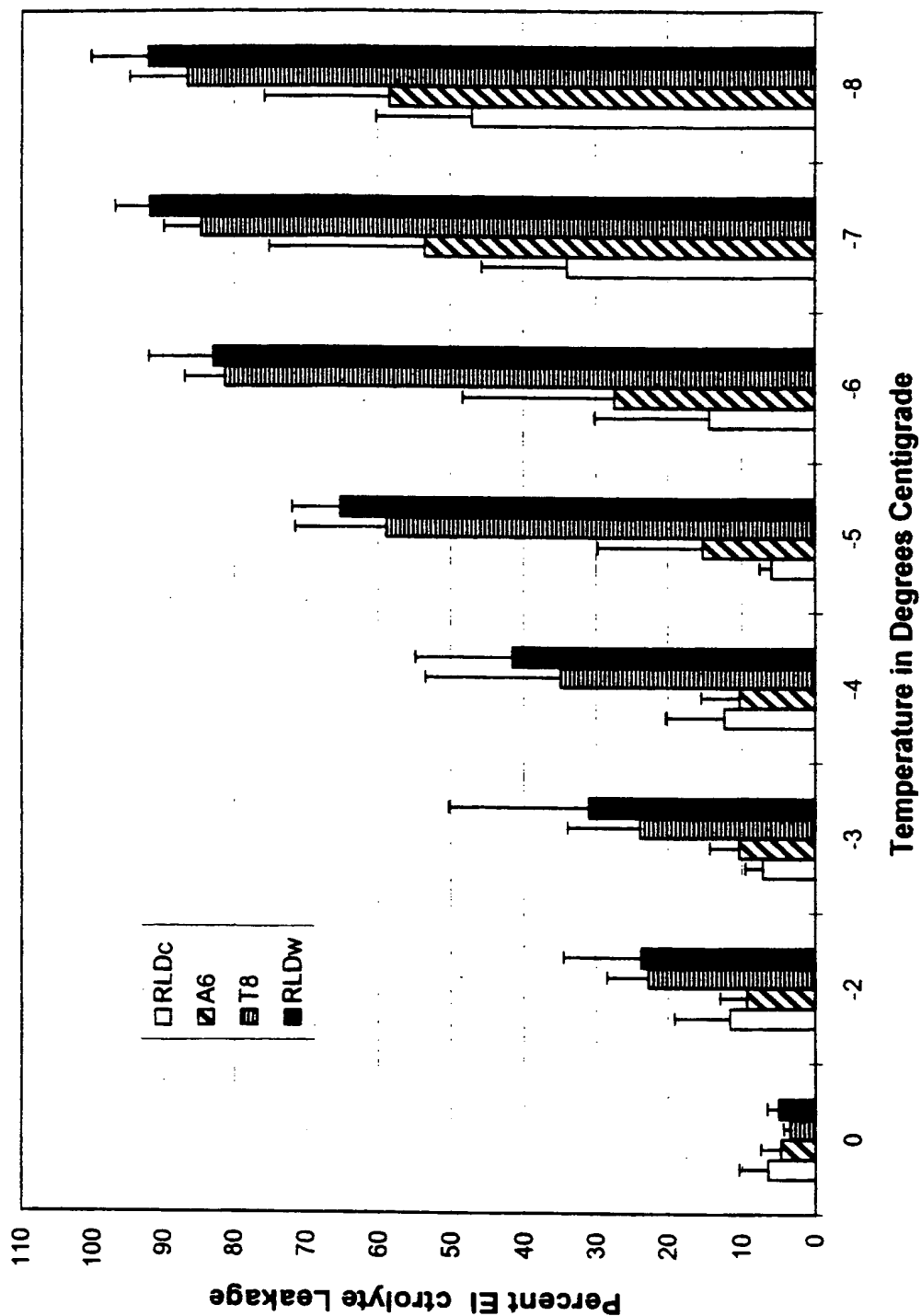


FIGURE 10B

SUBSTITUTE SHEET (RULE 26)

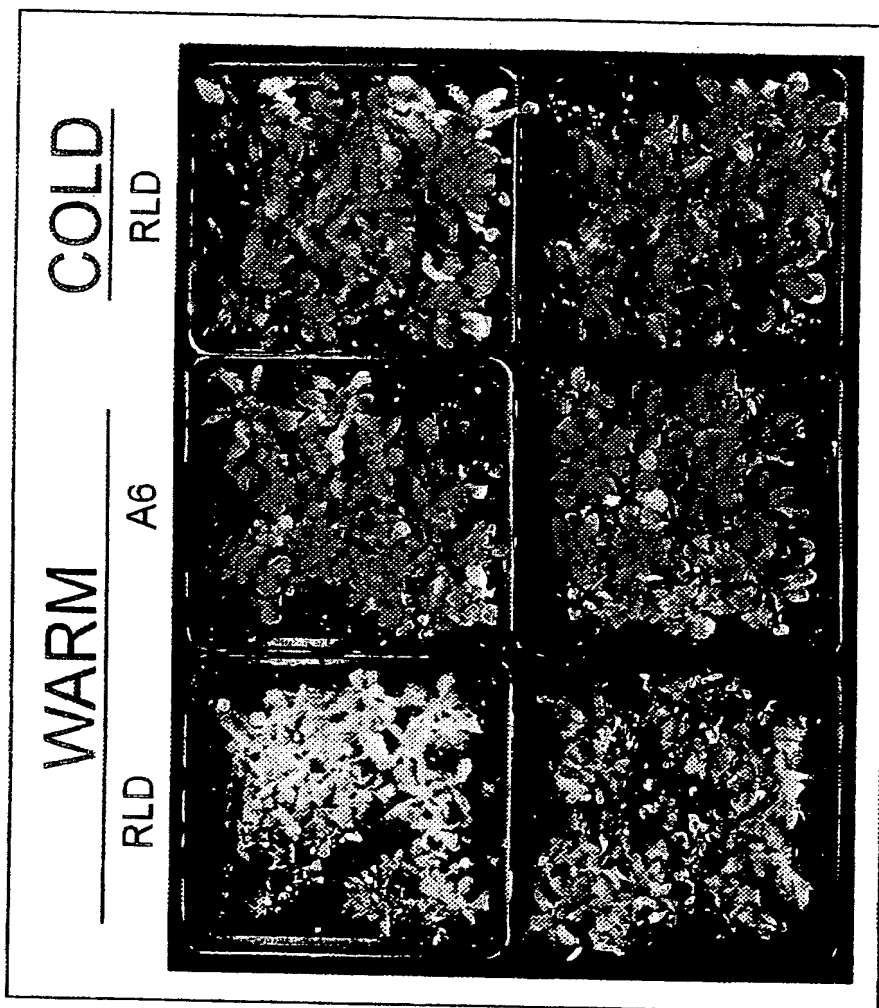


FIGURE 11

ATGAACTCATTTTCTGCCTTTTCTGAAATGTTTGGCTCCGATTACGAGTCTCCGGTTTCC 60  
 Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr Glu Ser Pro Val Ser  
 TCAGGCGGTGATTACAGTCCGAAGCTTGGCCACGAGCTGCCCAAGAAACACGCGGAAGG 120  
 Ser Gly Gly Asp Tyr Ser Pro Lys Leu Ala Thr Ser Cys Pro Lys Lys Pro Ala Gly Arg  
 AAGAAGTTTTCGTGAGACTCGTCACCCCAATTACAGAGGAGTTTCGTCAAAGAAACTCCGGT 180  
 Lys Lys Phe Arg Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly  
 AAGTGGGTGTGTGAGTTGAGAGAGCCCAACAAGAAACGAGGATTTGGCTCGGGACTTTC 240  
 Lys Trp Val Cys Glu Leu Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly Thr Phe  
 CAACCCGCTGAGATGGCAGCTCGTGTCTCAGCAGCTCGCCGCCATAGCTCTCCGTGGCAGA 300  
 Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala Ala Ile Ala Leu Arg Gly Arg  
 TCTGCCTGTCTCAATTTTCGTGACTCGGCTTGGCGGTACGAATCCCGGAATCAACCTGT 360  
 Ser Ala Cys Leu Asn Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys  
 GCCAAGGAAATCCA'AAAGGCGGCGGTGAAGCCGCTTGAATTTTCAAGATGAGATGTGT 420  
 Ala Lys Glu Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Asn Phe Gln Asp Glu Met Cys  
 CATATGACGACGGATGCTCATGGTCTTGACATGGAGGAGACCTTGGTGGAGGCTATTAT 480  
 His Met Thr Thr Asp Ala His Gly Leu Asp Met Glu Thr Leu Val Glu Ala Ile Tyr  
 ACGCCCGGAACAGAGCCCAAGATGCGTTTATATGGATGAAGAGGCGGATGTGGGGATGTCT 540  
 Thr Pro Glu Gln Ser Gln Asp Ala Phe Tyr Met Asp Glu Glu Ala Met Leu Gly Met Ser  
 AGTTTGTGGATAACATGGCCGAAGGGATGCTTTTACCGTCGCCGTCGGTTCATGAATGGAAC 600  
 Ser Leu Leu Asp Asn Met Ala Glu Gly Met Leu Leu Pro Ser Pro Ser Val Gln Trp Asn  
 TATAATTTTGTGTCGAGGGAGATGATGACGTGTCTTATGGAGCTATTAA 651  
 Tyr Asn Phe Asp Val Glu Gly Asp Asp Val Ser Leu Trp Ser Tyr •

FIGURE 12

SUBSTITUTE SHEET (RULE 26)

ATGAACATCTTCTGCTTTTCTGAAATGTTGGCTCCGATTACGAGTCTTCGGTTTCC 60  
 Met Asn Ser Phe Ser Ala Phe Ser Glu Met Phe Gly Ser Asp Tyr Glu Ser Ser Val Ser  
 TCAGCGGTGATTATATCCGACGCTTGCGAGCAGCTGCCCCCAAGAAACCGCGGTCGT 120  
 Ser Gly Gly Asp Tyr Ile Pro Thr Leu Ala Ser Ser Cys Pro Lys Lys Pro Ala Gly Arg  
 AAGAAGTTTCGTGAGACTCGTCACCCCAATATACAGAGGAGTTCGTCGGAGAAACTCCGGT 180  
 Lys Lys Phe Arg Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Arg Asn Ser Gly  
 AAGTGGGTTTGTGAGGTTAGAGAACCAACAAAGAAACAAAGGATTTGGCTCGGAACATTT 240  
 Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu Gly Thr Phe  
 CAAACCGCTGAGATGGCAGCTCGAGCTCAGCAGCTTGCCGCTTAGCCCTTCGTGGCCGA 300  
 Gln Thr Ala Glu Met Ala Ala Arg Ala His Asp Val Ala Leu Ala Leu Arg Gly Arg  
 TCAGCCTGCTCAATTTGCTGACTCGGCTTGAGACTCCGAATCCCGGAATCAACTTGC 360  
 Ser Ala Cys Leu Asn Phe Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Ser Thr Cys  
 GCTAAGGACATCCAAAAGGCGCGCTGAAGCTGCGTTGGCGTTTCAGGATGAGATGTGT 420  
 Ala Lys Asp Ile Gln Lys Ala Ala Glu Ala Leu Ala Phe Gln Asp Glu Met Cys  
 GATGCGACGCGATCATGGCTTCGACATGGAGGAGACGTTGGTGGAGGCTATTTACACG 480  
 Asp Ala Thr Thr Asp His Gly Phe Asp Met Glu Glu Thr Leu Val Glu Ala Ile Tyr Thr  
 GCGGAACAGAGCGAAATGCGTTTATATGCACGATGAGGCGATGTTTGAGATGCCGAGT 540  
 Ala Glu Gln Ser Glu Asn Ala Phe Tyr Met His Asp Glu Ala Met Phe Glu Met Pro Ser  
 TTGTTGGCTAATATGGCAGAAGGGATGCTTTTGCCGCTTCCGTCCGTACAGTGGGAATCAT 600  
 Leu Leu Ala Asn Met Ala Glu Gly Met Leu Leu Pro Leu Pro Ser Val Gln Trp Asn His  
 AATCATGAAGTCGACGCGGATGATGACGACGATCGTTATGGAGTTATTAA 651  
 Asn His Glu Val Asp Gly Asp Asp Val Ser Leu Trp Ser Tyr •

FIGURE 13

SUBSTITUTE SHEET (RULE 26)

cbf1.pro	MNSFSAFSEMGSDYESXVSSGGDYXPTLATSCPKKPA	10	20	30	40	50
cbf2.PRO	MNSFSAFSEMGSDYEPQ--GGDYCPTLATSCPKKPA					50
cbf3.PRO	MNSFSAFSEMGSDYESPVSSGGDYSPKLATSCPKKPA					50
	MNSFSAFSEMGSDYESSVSSGGDYIPTLASSCPKPA					50
	YRGVQRNSGKVVCEVREPNNKKTIRIWLGTFTQTA	60	70	80	90	100
cbf1.pro	YRGVQRNSGKVVSEVREPNNKKTIRIWLGTFTQTA					100
cbf2.PRO	YRGVQRNSGKVVCELREPNNKKTIRIWLGTFTQTA					100
cbf3.PRO	YRGVRRNSGKVVCEVREPNNKKTIRIWLGTFTQTA					100
	SACLNFAADSARLRIPESTCAKDIQKAAEAALAFQ	110	120	130	140	150
cbf1.pro	SACLNFAADSARLRIPESTCAKDIQKAAEAALAFQ					150
cbf2.PRO	SACLNFAADSARLRIPESTCAKEIQKAAEAALNFQ					150
cbf3.PRO	SACLNFAADSARLRIPESTCAKDIQKAAEAALAFQ					150
	MEETLVEAIYTPQSEXAFYMDEEAMFGMPSLLDN	160	170	180	190	200
cbf1.pro	MEETLVEAIYTPQSEXAFYMDEEAMFGMPSLLDN					200
cbf2.PRO	MEETLVEAIYTPQSQDAFYMDDEEAMLGSSLLDN					200
cbf3.PRO	MEETLVEAIYTAEQSENAFYMHDEAMFEMPSLLAN					200
	HNXDVEGDDDD-VSLWSY					
	HNXDVEGDDDD-VSLWSY					
cbf1.pro	HNXDVEGDDDD-VSLWSY					210
cbf2.PRO	HNXDVEGDDDD-VSLWSY					210
cbf3.PRO	HNXDVEGDDDD-VSLWSY					210

**FIGURE 14**

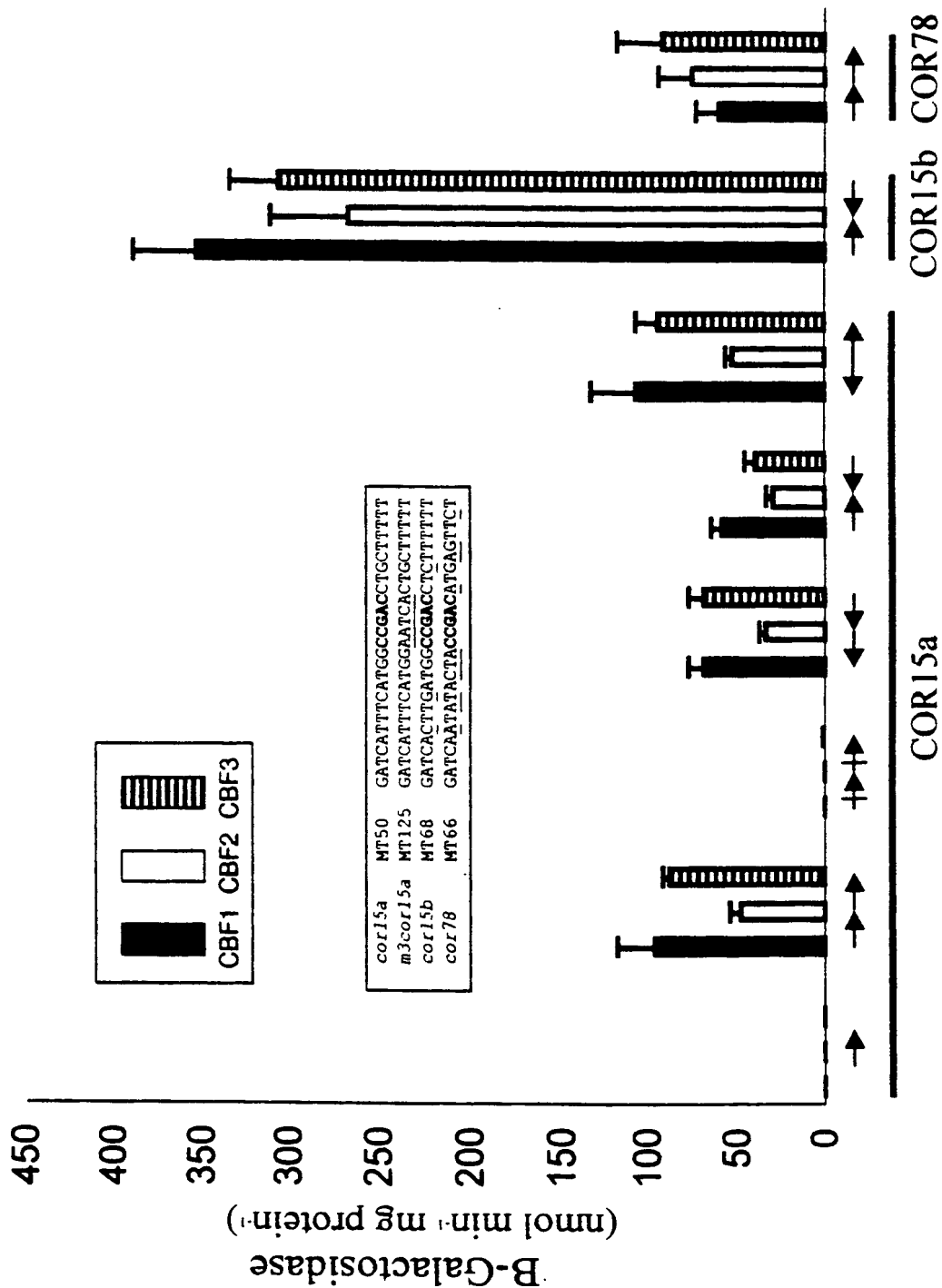


FIGURE 15

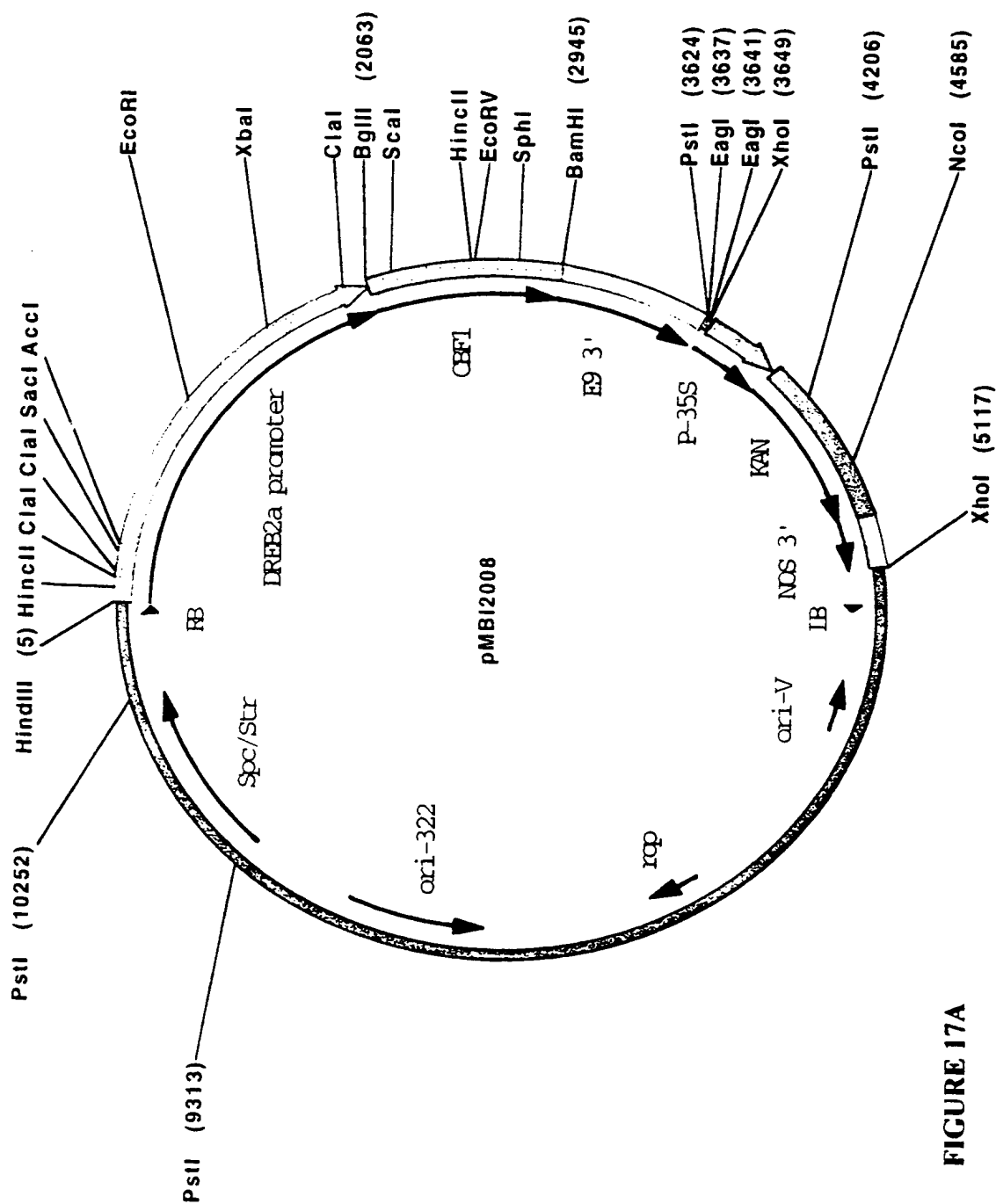


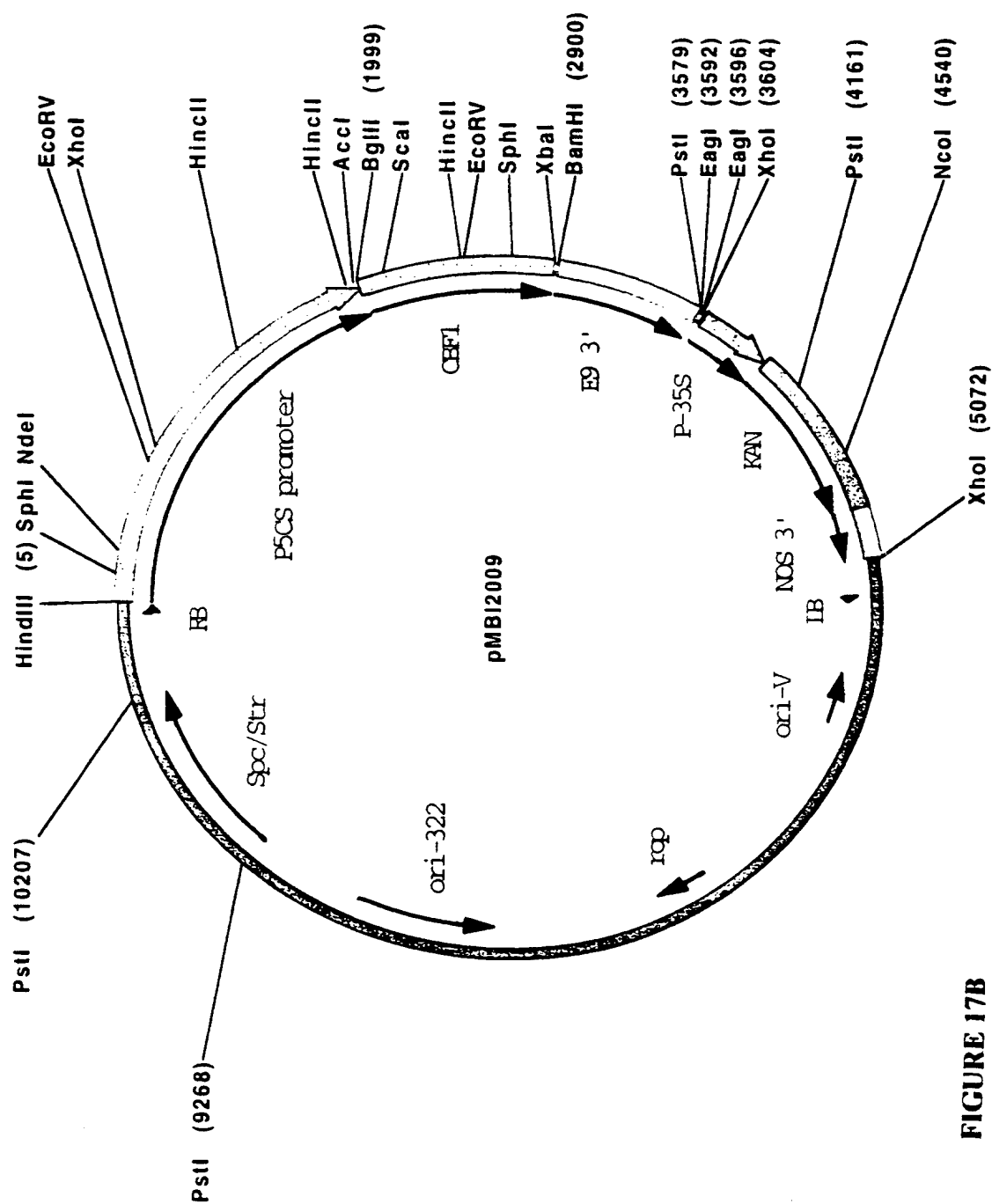
```

1  mnsefafeem fgadyepggg dycptlatac pkkpagrkkf retrHPIYRG 50
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus
51  VRQRNSGKMW SEVREPNNKt RIWLGTFqTA EMAARAHdVA ALALRGRsAC 100
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus
101  LNfADSAWRL RIPEsTCaKD IQKAAAEaAL AFq.....dET 150
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus
151  c.....DT.....TTTDHG lDMEETMVEA 200
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus
201  iYtPEQSEG. ....a fYNDEEtMfg MPTLLdnMAE 250
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus
251  GMLLpppsvg wnhnydgegd gdvsIwEY 278
   at-nap{cbf1}
   at-nap{napus-homolog}
   Consensus

```

FIGURE 16





**FIGURE 17B**

**SUBSTITUTE SHEET (RULE 26)**

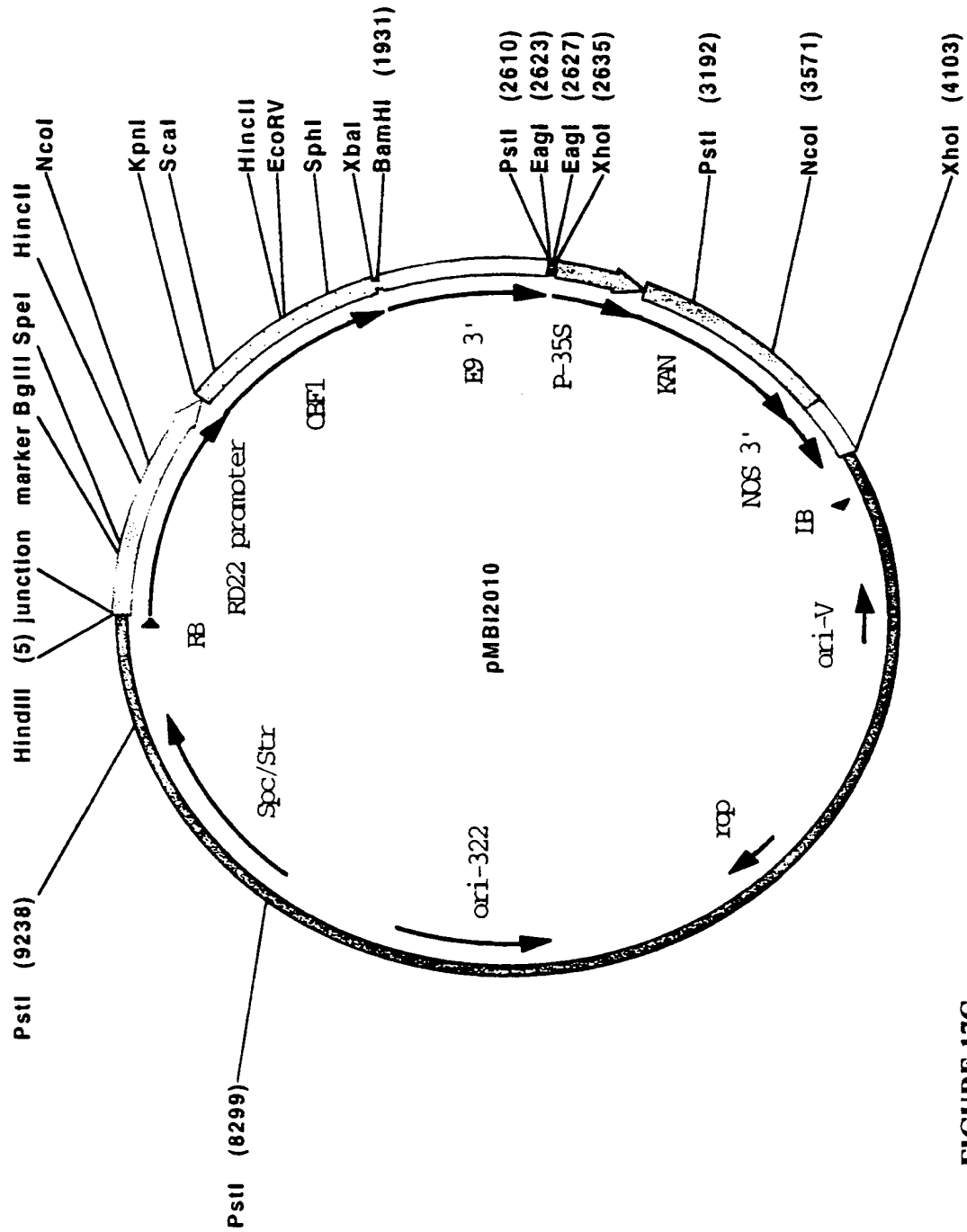
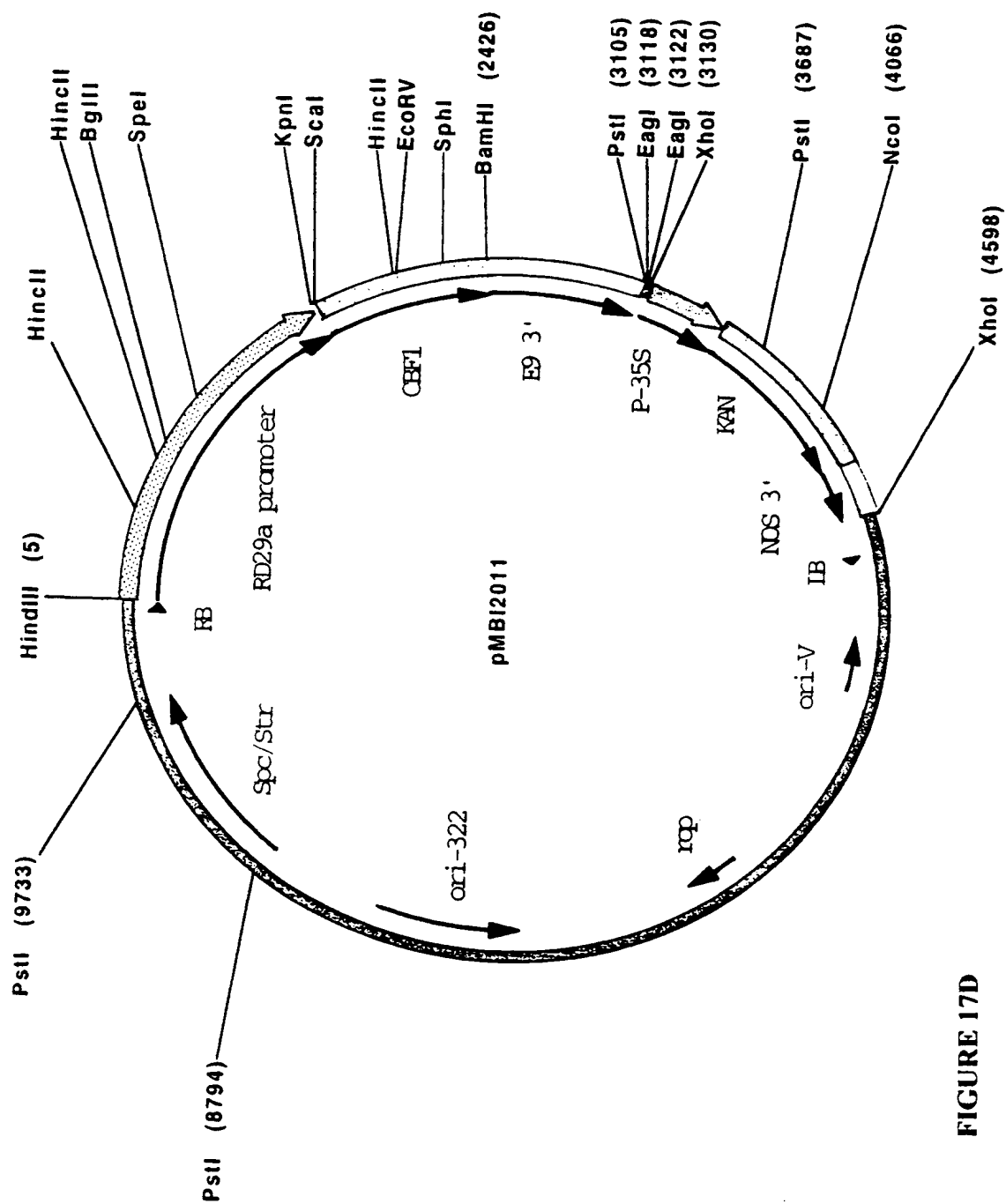


FIGURE 17C



**FIGURE 17D**

**SUBSTITUTE SHEET (RULE 26)**

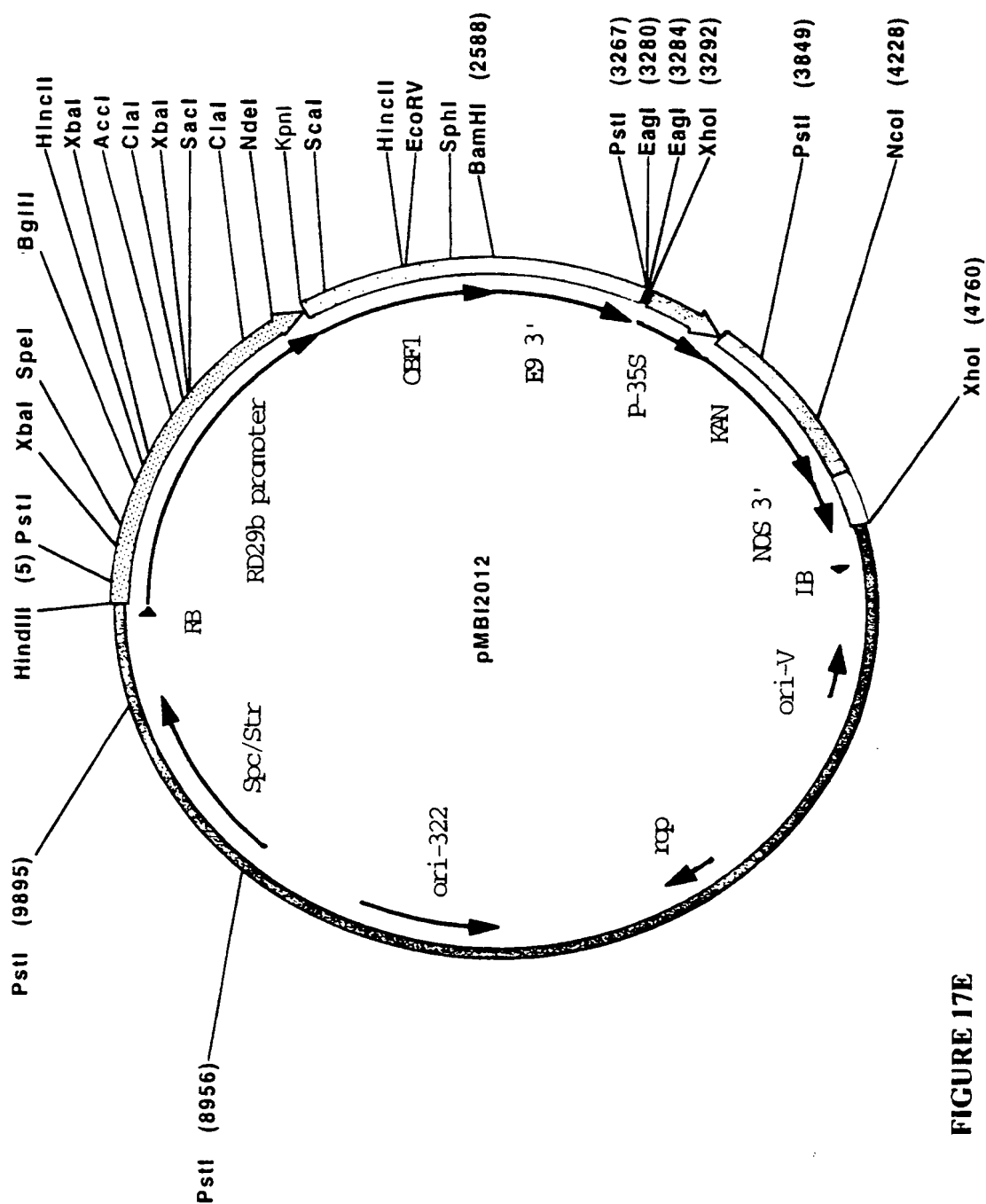


FIGURE 17E

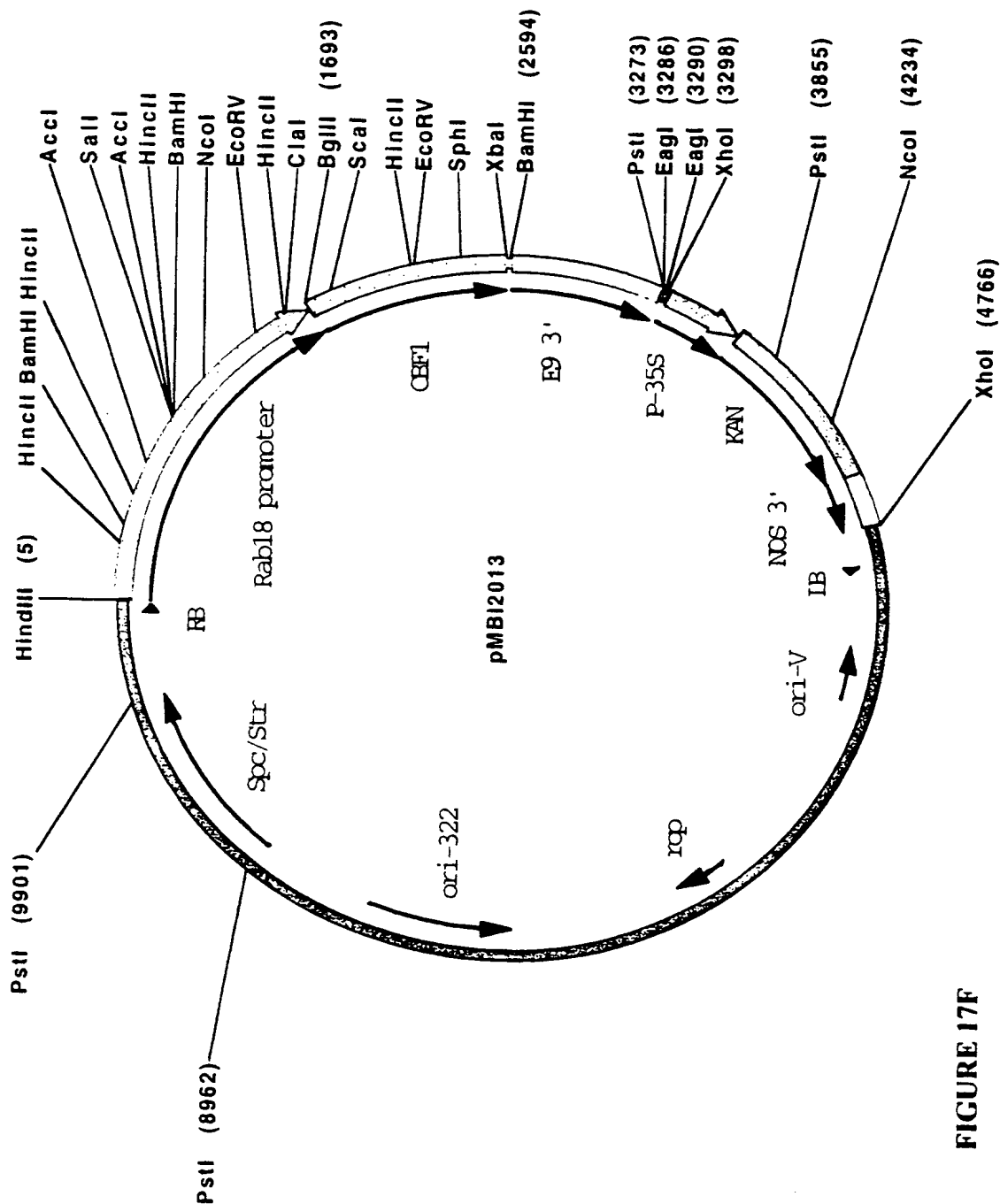


FIGURE 17F

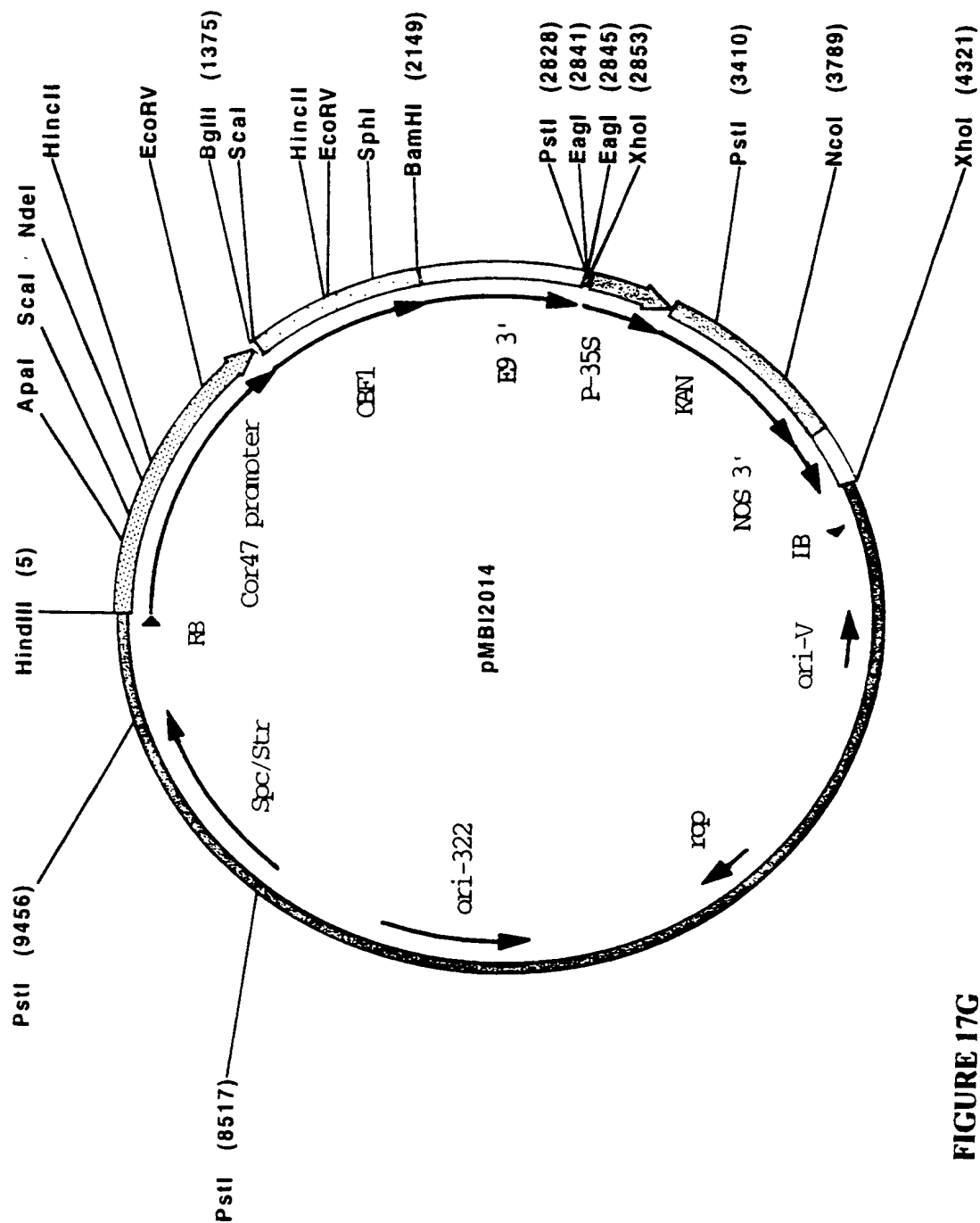


FIGURE 17C



27/45

**FIGURE 18A**

**bjCBF1** Species=Brassica juncea Length=577 [SEQ ID No. 38]  
TTTCACCCTATCTACCGGGGAGTTTCGCTGAGAAAAGTCAGGTAAGTGGGTGTGTGAAGTG  
AGGGAGCCAAACAAGAAATCTAGGATTTGGCTTGGAACTTTCAAACCCGAGAGATCGCT  
GCTCGTGCTCAGGACGTTGCCGCTTAGCCCTCCGTGGAAGAGCGGCCTGTCTCAACTTC  
GCCGACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACTTGCGCCAAGGATATCCAGAAG  
GCTGCTGCTGAAGCTGCGTTGGCTTTTGGGGCCGAAAAGAGTGATACCACGACGAATGAT  
CAAGGCATGAACATGGAGGAGATGACGGTGGTGGCTTCTCAGGCTGAGGTGAGCGACACG  
ACGACATATCATGGCCTGGACATGGAGGAGACTATGGTGGAGGCTGTTTTTGGCTGAGGAA  
CAGAGAGAAGGGTTTTACTTGGCGGAGGAGACGACGGTGGAGGGTGTGTTTACGGAGGAA  
CAGAGCAAAGGGTTTTATATGTACGAGGAGTGGACGTTCCGGATGCAGTCCTTTTTGGCC  
GATATGGCTGAAGGCATGCTCTTTTCAAAGGGCGAAT

**bjCBF2** Species=Brassica juncea Length=431 [SEQ ID No. 40]  
CATCCGATCTACAGGGGAGTTTCGTCTGAGAAAATCAGGTAAGTGGGTGTGTGAAGTGAGG  
GAACCAAACAAGAGATCTAGGATCTGGCTCGGTACTTTCTTAACCGCCGAGATCGCAGCT  
CGCGCTCAGGACGTCGCCGCCATAGCCCTCCGTGGCAAATCCGCATGTCTCAATTTTCGCT  
GACTCGGCTTGGCGGCTCCGTATCTCGGAGACAAACATGCCCTAAGGAGATTGAGAAGGCT  
GCTGCTGAAGCCCGGCTGGCTTTTTCAGGCTGAGCTAAATGATACGACGGCCGATCATGGC  
CTTGACGTGGAGGAGACGATCGTGGAGGCTATTTTTCACGGAGGAAAGCAGCGAAGGGTTT  
TATATGGACGAGGAGTTCATGTTCCGGATGCCGACCTTGTGGGCTAGTATGGCAGAAGGG  
ATGCTTCTTCC

**bjCBF3** Species=Brassica juncea Length=431 [SEQ ID No. 42]  
CATCCAATTTACCGTGGAGTTTCGTCTGAGAAAATCAGGTAAGTGGGTGTGTGAAGTGAGG  
GAGCCAAACAAGAAATCTAGGATCTGGCTCGGTACTTTCTTAACCGCCGAGATCGCAGCT  
CGCGCTCAGGACGTCGCCGCCATAGCCCTCCGTGGCAAATCCGCATGTCTCAATTTTCGCT  
GACTCGGCTTGGCGGCTCCGTATCCCGGAGACAAACATGCCCTAAGGAGATTGAGAAGGCT  
GCTGCTGAAGCCCGGCTGGCTTTTTCAGGCTGAGCTAAATGATACGACGGCCGATCATGGC  
CTTGACGTGGAGGAGACGATCGTGGAGGCTATTTTTCACGGAGGAAAGCAGCGAAGGGTTT  
TATATGGACGAGGAGTTCATGTTCCGGATGCCGACCTTGTGGGCTAGTATGGCGGAGGGC  
ATGCTCCTTCC

**bjCBF4** Species=Brassica juncea Length=425 [SEQ ID No. 44]  
CATCCAATCTACCGGGGTGTTTCGACAGAGAAACTCAGGGAATGGGTTTGTGAAGTTAGG  
GAGCCTAATAAGAAATCTAGGATCTGGTTAGGGACTTTTCCGACCGTCGAAATGGCCGCT  
CGTGCTCAGGACGTCGCCGCTTTAGCCCTTCGTGGCCGCTCCGCTTGTCTTAATTTTCGCC  
GACTCGGCGTGGTGTCTACGGATTCCCGAGTCTACTTGTCTTAAAGAGATTGAGAAAGCT  
GCCGCCGAAGCTGCAATGGCGTTTTCAGAACGAGACGGCTACGACTGAGACGACTATGTTT  
GAGGGAGTCAATACCGGCGGAGGAGACGGTGGGGCAGACGCGTGTGGAGACAGCAGAGGAG  
AACGGTGTGTTTTATATGGACGATCCAAGGTTTCTTGAGAATATGGCAGAGGGCATGTTCT  
TACC

**bnCBF1** Species=Brassica napus Length=632 [SEQ ID No. 46]  
CACCCGATATACCGGGGAGTTTCGTCTGAGAAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGG  
GAACCAAACAAGAAATCTAGAATTTGGCTTGGAACTTTCAAACAGCTGAGATGGCAGCT  
CGTGCTCAGGACGTCGCTGCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCTCAATTATGCG  
GACTCGGCTTGGCGGCTCCGCATCCCGGAGACAACTGCCACAAGGATATCCAGAAGGCT  
GCTGCTGAAGCCGCTTGGCTTTTTCAGGCTGAGAAAAGTGATGTGACGATGCAAAATGGC  
CAGAACATGGAGGAGACGACGGCGGTGGCTTCTCAGGCTGAAGTGAATGACACGACGACA  
GAACATGGCATGAACATGGAGGAGGCAACGGCAGTGGCTTCTCAGGCTGAGGTGAATGAC  
ACGACGACGGATCATGGCGTAGACATGGAGGAGACAATGGTGGAGGCTGTTTTTACTGGG  
GAACAAAGTGAAGGGTTTAAACATGGCGAAGGAGTCCGACGGTGGAGGCTGCTGTTGTACG  
GAGGAACCGAGCAAAGGATCTTACATGGACGAGGAGTGGATGCTCGAGATGCCGACCTTG  
TTGGCTGATATGGCAGAAGGGATGCTCCTGCC

**bnCBF2** Species=Brassica napus Length=876 [SEQ ID No. 48]  
ACCGCTCGAGCAACAATGAACACATTCCCTGCTTCCACTGAAATGGTTGGCTCCGAGAAC  
GAGTCTCCGTTACTACGGTAGTAGGAGGTGATTATTATCCCATGTTGGCGGCAAGCTGT  
CCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGAGACACGTCACCCCATTTACCGAGGA  
GTTCTGCTGAGAAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAACAAGAAATC

**SUBSTITUTE SHEET (RULE 26)**

## FIGURE 18A-continued

AGAATTTGGCCCCGGAACCTTTCAAAACAGCTGAGATGGCAGCTCGTGCTCACGACGTCGCT  
GCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCTCAATTATGCGGACTCGGCTTGGCGGCTC  
CGCATCCCGGAAACAACCTGCCACAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATTG  
GCTTTTGAGGCTGAGAAAAGTGATGTGACGATGCAAAATGGCCTGAACATGGAGGAGACG  
ACGGCGGTGGCTTCTCAGGCTGAAGTGAATGACACGACGACAGAACATGGCATGAACATG  
GAGGAGGCAACAGCGGTGGCTTCTCAGGCTGAGGTGAATGACACGACGACAGATCATGGC  
GTAGACATGGAGGAGACGATGGTGGAGGCTGTTTTTACGGAGGAACAAAGTGAAGGGTTC  
AACATGGCGGAGGAGTCGACGGTGGAGGCTGCTGTTGTTACGGATGAACTGAGCAAAGGA  
TTTTACATGGACGAGGAGTGGACGTACGAGATGCCGACCTTGTGGCTGATATGGCGGCA  
GGGATGCTTTTGGCGCCACCATCTGTACAATGGGGACATAATGATGACTTGAAGGAGAT  
GCGGACATGAACCTCTGGAGTTATTAAGGATCCGCG

**bnCBF3** Species=Brassica napus Length=884 [SEQ ID No. 50]  
ACTACACTCAGCCTTATCCAGTTTTTTTCAAAAGATTTTTCAACAATGAACACATTCCT  
GCGTCCACTGAAATGGTTGGCTCCGAGAACGAGTCTCCGGTTACTACGGTAGCAGGAGGT  
GATTATTATCCCATGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCAGGTAGGAAGAAGTTT  
CAGGAGACACGTACCCCCATTTACCGAGGAGTTCGTCTGAGAAAAGTCAGGTAAAGTGGTG  
TGTGAAGTGAGGGAACCAACAAGAAATCTAGAATTTGGCCCCGGAACCTTTCAAAACAGCT  
GAGATGGCAGCTCGTGCTCACGACGTGCTGCCCTAGCCCTCCGTGGAAGAGGCGCCTGC  
CTCAATTATGCGGACTCGGCTTGGCGGCTCCGCATCCCGGAGACAACCTGCCACAAGGAT  
ATCCAGAAGGCTGCTGCTGAAGCCGCATTGGCTTTTGAGGCTGAGAAAAGTGATGTGACG  
ATGCAAAATGGCCTGAACATGGAGGAGACGACGGCGGTGGCTTCTCAGGCTGAAGTGAAT  
GACACGACGACAGAACATGGCATGAACATGGAGGAGGCAACGGCAGTGGCTTCTCAGGCT  
GAGGTGAATGACACGACGACGGATCATGGCGTAGACATGGAGGAGACAATGGTGGAGGCT  
GTTTTTACTGGGGAACAAAGTGAAGGGTTTAACTGGCGAAGGAGTCGACGGTGGAGGCT  
GCTGTTGTTACGGAGGAACCGAGCAAAGGATCTTACATGGACGAGGAGTGGATGCTCGAG  
ATGCCGACCTTGTGGCTGATATGGCGGAAGGGATGCTTTTGGCGCCGCGCTCCGTACAA  
TGGGGACAGAATGATGACTTCGAAGGAGATGCTGACATGAACCT

**bnCBF4** Species=Brassica napus Length=874 [SEQ ID No. 52]  
GTAATTCGATTACCGCTCGAGTACTTACTATACTACACTCAGCCTTATCCAGTTTTTCAA  
AAGAAGTTTTCAACTATGAACTCAGTCTCTACTTTTTCTGAACTTCTTGGCTCTGAGAAC  
GAGTCTCCGGTAGGTGGTGATTACTGTCCCATGTTGGCGGCGAGCTGTCCGAAGAAGCCG  
GCGGGTAGGAAGAAGTTTCCGGAGACACGTACCCCCATTTACCGAGGAGTTCGCTTAGA  
AAATCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAAACAAAAATCTAGGATTTGGCTC  
GGAACCTTTCAAAACAGCTGAGATCGCAGCTCGTGCTCACGACGTGCGCGCTTAGCTCTC  
CGTGGAAGAGGCGCCTGCCTCAACTTCGCCGACTCGGCTTGGCGGCTCCGTATCCCGGAG  
ACAACCTGCGCCAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATTGGCTTTTGAGGCC  
GAGAAGAGTGATACCACGACGAATGATCATGGCATGAACATGGCTTCTCAGGCCGAGGT  
AATGACACAACGGATCATGGCCTGGACATGGAGGAGACGATGGTGGAGGCTGTTTTTACT  
GAGGAGCAGAGAGACGGTTTTTACATGGCGGAGGAGACGACGGTGGAGGGTGTGTTCCG  
GAGGAACAGATGAGCAAAGGGTTTTACATGGACGAGGAGTGGATGTTCCGGATGCCGACC  
TTGTTGGCTGATATGGCGGCAGGGATGCTCTTACCGCCGCGCTCCGTACAATGGGGACAT  
AATGATGACTTCGAAGGAGATGTTGACATGAACCTCTGGAATTATTAGTACTCATATTTT  
TTTAAATTATTTTTTGAACGAATAATATTTTATT

**bnCBF1** Species=Brassica napus Length=898 [SEQ ID No. 54]  
AATAAATATCTTATCAAACCAGTCAGAACAGAGATCTTGTTACTTACTATACTACACTCA  
GCCTTATCCAGTTTTTCAAAAAAGTATTCAACGATGAACTCAGTCTCTACTTTTTCTGAA  
CTGCTCCGCTCCGAGAACGAGTCTCCGGTTAATACGGAAGGTGGTGATTACATTTTGGCG  
GCGAGCTGTCCCAAGAAACCTGCTGGTAGGAAGAAGTTTCAGGAGACACGCCACCCCAT  
TACAGAGGAGTTCGTCTGAGGAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAC  
AAGAAATCTAGAATTTGGCTCGGAACCTTTCAAAACAGCTGAGATCGCAGCTCGTGCTCAC  
GACGTTGCCGCCTTAGCTCTCCGTGGAAGAGGCGCCTGCCTCAACTTCGCCGACTCGGCT  
TGGCGGCTCCGTATCCCGGAGACGACCTGCGCCAAGGATATCCAGAAGGCTGCTGCTGAA  
GCCGATTGGCTTTTGAGGCCGAGAAGAGTGATACCACGACGAATGATCATGGCATGAAC  
ATGGCTTCTCAGGTTGAGGTTAATGACACGACGGATCATGACCTGGACATGGAGGAGACG  
ATAGTGGAGGCTGTTTTTAGGGAGGAACAGAGAGAAGGGTTTTACATGGCGGAGGAGACG  
ACGGTTGTGGGTGTGTTCCGGAGGAACAGATGAGCAAAGGGTTTTACATGGACGAGGAG  
TGGATGTTCCGGGATGCCGACCTTGTGGCTGATATGGCGGCAGGGATGCTCTTACCGCTG  
CCGTCCGTACAATGGGGACATAATGATGACTTCGAAGGAGATGCTGACATGAACCTCTGG  
AATTATTAGTACTCATATTTTTTTAAATTATTTTTTGAACGAATAATATTTTATTGAA

**bnCBF6** Species=Brassica napus Length=1132 [SEQ ID No. 56]  
GATTACCGCTCGAGTACTTACTATACTACACTCAGCCTTATCCAGTTTTTCTCAAAAGAT  
TTTTCAACAATGAACACATTCCTGCTTCCACTGAAATGGTTGGCTCCGAGAACGAGTCT

SUBSTITUTE SHEET (RULE 26)

## FIGURE 18A-continued

CCGGTTACTACGGTAGTAGGAGGTGATTATTATCCCATGTTGGCGGCAAGCTGTCCGAAG  
AAGCCAGCGGGTAGGAAGAAGTTTTCAGGAGACACGTACCCCCATTTACCGAGGAGTTTCGT  
CTGAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAACAAGAAATCTAGAATT  
TGGCTTGGAACCTTTCAAACAGCTGAGATGGCAGCTCGTGCTCACGACGTGGCTGCCCTA  
GCCCTCCGTGGAAGAGGCGCCTGCCTCAATTATGCGGACTCGGCTTCGCGGCTCCGCATC  
CCGGAGACAACCTGCCACAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATTGGCTTTT  
GAGGCTGAGAAAAGTGATGTGACGATGGAGGAGACGATGGCGGTGGCTTCTCAGGCTGAA  
GTGAATGACACGACGACAGATCATGGCATGAACATGGAGGAGGCAACAGCGGTGGCTTCT  
CAGGCTGAGGTGAATGACACGACGACAGATCATGGCGTAGACATGGAGGAGACGATGGTG  
GAGGCTGTTTTTACGGAGGAACAAAGTGAAGGGTTCAACATGGCGGAGGAGTCGACGGTG  
GAGGCTGCTGTTGTTACGGATGAACTGAGCAAAGGATTTTACATGGACGAGGAGTGGACG  
TACGAGATGCCGACCTTGTGGCTGATATGGCGGCAGGGATGCTTTTGCCGCCACCATCT  
GTACAATGGGGACATAATGATGACTTGGAAAGGAGATGCTGACATGAACCTCTGGAATTAT  
TAATACTCGTGTTTTAAATTTATACATTGTGCAATAATTTTTATCGAATTTCTAATTC  
TGCCTTTAACTTTAATGGGGATCTTTATTAGTGTAGGAAACGAGTGTAAATGTTCCGCC  
GTGGTGTGTCAAAATGCTGATTATTTTTTGTGTGACGACATAATCACGTTTGGTTTCCTT  
TACACTCCAAATTTAGTTGAAATACAAATAGAATAGAAAAGTGAAAAATGT

**bnCBF7** Species=Brassica napus Length=768 [SEQ ID No. 58]  
AGTGATGTTTTTCAAAGAAGTTTTCAACTATGAACTCAGTCTCTACTTTTTCTGAACTT  
CTTGGCTCTGAGAACGAGTCTCCGGTAGGTGGTGATTACTGTCCCATGTTGGCGGCGGAGC  
TGTCCGAAGAAGCCGGCGGGTAGGAAGAAGTTTCGGGAGACACGTACCCCCATTTACCGA  
GGAGTTTCGCTTAGAAAATCAGGTAAGTGGGTGTGTGAAGTGAGGGAGCCAAACAAGAAA  
TCTAGGATTTGGCTCGGTACTTTCTAACAGCCGAGATCGCAGCCCGTGCTCACGACGTC  
GCCGCATAGCCCTCCGCGGCAAAATCAGCTTGTCTCAATTTTGCCGACTCCGCTTGGCGG  
CTCCGTATCCCGGAGACAACATGCCCCAAGGAGATTGAGAAGCGGCTGCTGAAGCCGCG  
GTGGCTTTTAAAGGCTGAGATAAATAACGACGGCGGATCATGGCATTGACGTGGAGGAG  
ACGATCGTTGAGGCTATTTTACGGAGGAAAACAACGATGGTTTTTATATGGACGAGGAG  
GAGTCCATGTTTCGGGATGCCGCCCTTGTGGCTAGTATGGCTGAAGGAATGCTTTTGCCG  
CCTCCGTCCGTACAATTCGGACATACCTATGACTTTGACGGAGATGCTGACGTGTCCCTT  
TGGAGTTATTAGTACAAAGATTTTTTATTTCCATTTTTTGGTATAATACTTCTTTTGATT  
TTCCGATTCTACCTTTTTTATGGGTATCATTTTTTTTTTTAGGAAACGGG

**bnCBF8** Species=Brassica napus Length=953 [SEQ ID No. 60]  
ACCGCTCGAGCAACAATGAACACATTCCCTGCTTCCACTGAAATGGTTGGCTCCGAGAAC  
GAGTCTCCGGTTACTACGGTAGCAGGAGGTGATTATTATCCCATGTTGGCGGCAAGCTGT  
CCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGAGACACGTACCCCCATTTACCGAGGA  
GTTTCGTCTGAGAAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAACAAGAAATCT  
AGAATTTGGCTTGGAACCTTTCAAACAGCTGAGATGGCAGCTCGTGCTCACGACGTGGCT  
GCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCTCAATTATGCGGACTCGGCTTCGCGGCTC  
CGCATCCCGGAGACAACCTGCCACAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATTG  
GCTTTTGAGGCTGAGAAAAGTGATGTGACGATGGAGGAGACGATGGCGGTGGCTTCTCAG  
GCTGAAGTGAATGACACGACGACAGATCATGGCATGAACATGGAGGAGGCAACGGCAGTG  
GCTTCTCAGGCTGAGGTGAATGACACGACGACGGATCATGGCGTAGACATGGAGGAGACA  
ATGGTGGAGGCTGTTTTTACTGGGGAACAAAGTGAAGGGTTTAAATATGGCGAAGGAGTCG  
ACGGTGGAGGCTGCTGTTTACGGAGGAACCGAGCAAAGGATCTTACATGGACGAGGAG  
TGGATGCTCGAGATGCCGACCTTGTGGCTGATATGGCGGAAGGGATGCTTTTGCCGCCG  
CCGTCCGTACAATGGGGACAGAATGATGACTTCAAGGAGATGCGGACATGAACCTCTGG  
AGTTATTAATACTCGTATTTTTTAAATTTATTTATGTGCAATAATTTTTTATCGAATTC  
GAATTCGCTTTAATTTTTAATGGGGATCTTTATTTGCCAAAAAAAAAAAAA

**bnCBF9** Species=Brassica napus Length=889 [SEQ ID No. 62]  
CTAGTGATTACCGCTCGAGCAACAATGAACACATTCCCTGCTTCCACTGAAATGGTTGGC  
TCCGAGAACGAGTCTCCGGTTACTACGGTAGCAGGAGGTGATTATTATCCCATGTTGGCG  
GCAAGCTGTCCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGAGACACGTACCCCCATT  
TACCGAGGAGTTCGTCTGAGAAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAAAC  
AAGAAATCTAGAATTTGGCCCGGAACCTTTCAAACAGCTGAGATGGCAGCTCGTGCTCAC  
GAGCTCGCTGCCCTAGCCCTCCGTGGAAGAGGCGCCCGCTCAATTATGCGGACTCAGCT  
TGGCGGCTCCGCATCCCGGAGACAACCTGCCACAAGGATATCCAGAAGGCTGCTGCTGAA  
GCCGCATTGGCTTTTGGAGGCTGAGAAAAGTGATGTGACGATGCAAAATGGCCTGAACATG  
GAGGAGACGACGGCGGTGGCTTCTCAGGCTGAAGTGAATGACACGACGACAGAACATGGC  
ATGAACATGGAGGAGGCAACGGCAGTGGCTTCTCAGGCTGAGGTGAATGACACGACGACG  
GATCATGGCGTAGACATGGAGGAGACAATGGTGGAGGCTGTTTTTACTGGGGAACAAAGT  
GAAGGGTTTAAATATGGCGAAGGAGTCCAGCGGTGGAGGCTGCTGTTTACGGAGGAACCG  
AGCAAAGGATCTTACATGGACGAGGAGTGGATGCTCGAGATGCCGACCTTGTGGCTGAT  
ATGGCGGAAGGGATGCTTTTGCCGCCGCCGTCCGTACAATGGGGACAGAATGATGACTTC

SUBSTITUTE SHEET (RULE 26)

## FIGURE 18A-continued

GAAGGAGATGCGCACATGAACCTCTGGAGTTATTAAGGATCCGCGAATC

**boCBF1** Species=Brassica oleracea Length=563 [SEQ ID No. 64]  
CACCCCTATCTACCGGGGAGTTTCGCTGAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGG  
GAGCCAAACAAGAAATCTAGGATTTGGCTTGGAACTTCAAAAACCGCAGAGATCGCTGCT  
CGTGCTCACGACGTTGCCGCTTAGCCCTCCGTGGAAGAGCGGCCTGTCTCAACTTCGCC  
GACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACCTGCGCCAAGGATATCCAGAAGGCT  
GCTGCTGAAGCTGCGTTGGCTTTTGGGGCCGAAAAGAGTGATACCACGACGAATGATCAA  
GGCATGAACATGGAGGAGATGACGGTGGTGGCTTCTCAGGCTGAGGTGAGCGACACGACG  
ACATATCATGGCCTGGACATGGAGGAGACTATGGTGGAGGCTGTTTTGCTGAGGAACAG  
AGAGAAGGGTTTTACTTGGCGGAGGAGACGACGGTGGAGGGTGTGTTACGGAGGAACAG  
AGCAAAGGGTTTTATATGGACGAGGAGTGGACGTTCCGGATGCAGTCCTTTTGGCCGAT  
ATGGCTGAAGGCATGCTCTTTCC

**boCBF2** Species=Brassica oleracea Length=533 [SEQ ID No. 66]  
GAAACATAGATCTTTGTACTTACTATACTTCACCTTATCCAGTTTTATTTTTTATTTAT  
AAAGAGTTTTCAACAATGACCTCATTTTCTACCTTTTCTGAAGTGTGGGCTCCGAGCAT  
GAGTCTCCGGTTACATTAGGCGAAGAGTATTGTCCGAAGCTGGCCGCAAGCTGTCCGAAG  
AAACCAGCCGGCCGGAAGAAGTTTCGAGAGACGCGTCACCCAGTTTACAGAGGAGTTTCGT  
CTGAGAACTCAGGTAAGTGGGTGTGTGAAGTGAGGGAGCCAAACAAGAAATCTAGGATT  
TGGCTCGGTACTTTCTTAACAGCCGAGATCGCAGCCCGTGTCTACGACGTCGCGCCGATA  
GCCCTCCGCGGCAAAATCAGCTTGTCTCAATTTTGGCGACTCCGCTTGGCGGCTCCGTATC  
CCGGAGACAACATGCCCAAGGAGATTGAGAAGGCGGCTGCTGAAGCCGCGGTGGCTTTT  
AAGGCTGAGATAAATAATACGACGCGGATCACGGCTCGACATGGAAGAGAC

**boCBF3** Species=Brassica oleracea Length=887 [SEQ ID No. 68]  
ACTCAGCCTTATCCAGTTTTTCTCAAAAGATTTTTCAACAATGAACACATTCCCTGCTTC  
CACTGAAATGGTTGGCTCCGAGAACGAGTCTCCGGTACTACGGTAGTAGGAGGTGATTA  
TTATCCCATGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGA  
GACACGTCACCCCATTTACCGAGGAGTTTCGTCTGAGAAAGTCAGGTAAGTGGGTGTGTGA  
AGTGAGGGAACCAACAAGAAATCTAGAATTTGGCTTGGAACTTTCAAAACAGCTGAGAT  
GGCAGCTCGTGCTCACGACGTGGCTGCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCTCAA  
TTATGCGGACTCGGCTTGGCGGCTCCGCATCCCGGAGACAACCTGCCACAAGGATATCCA  
GAAGGCTGCTGCTGAAGCCGATTGGCTTTTGAGGCTGAGAAAAGTGATGTGACGATGGA  
GGAGACGATGGCGGTGGCTTCTCAGGCTGAAGTGAATGACACGACGACAGATCATGGCAT  
GAACATGGAGGAGGCAACAGCGGTGGCTTCTCAGGCTGAGGTGAATGACACGACGACAGA  
TCATGGCGTAGACATGGAGGAGACGATGGTGGAGGCTGTTTTACGGAGGAACAAAGTGA  
AGGCTTCAACATGGCGGAGGAGTGGACGCGTGGAGGCTGCTGTTGTTACGGATGAAGAG  
CAAAGGATTTTACATGGACGAGGAGTGGACGTACGAGATGCCGACCTTGTGGCTGATAT  
GGCGGCAGGGATGCTTTTGGCGCCACCATCTGTACAATGGGACATAATGATGACTTGGA  
AGGAGATGCGGACATGAACCTCTGGAGTTATTAATACTCGTATTTTT

**boCBF4** Species=Brassica oleracea Length=950 [SEQ ID No. 70]  
CTGAAAAGAAGATAAAAGAGAGAGAAATAAATATCTTATCAAACCAGACAGAACAGAGAT  
CTTGTTACTTACTATACTACACTCAGCCTTATCCAGTTTTTCAAAAGAAGTTTTCAACTA  
TGAAGCTCAGTCTCTACTTTTTCTGAAGTCTTGGCTCTGAGAACGAGTCTCCGGTAGGTG  
GTGATTACTGTCCCATGTTGGCGGCGAGCTGTCCGAAGAAGCCGGCGGGTAGGAAGAAGT  
TTCGGGAGACACGTCACCCCATTTACCGAGGAGTTCGCCTTAGAAAAATCAGGTAAGTGGG  
TGTGTGAAGTGAGGGAACCAACAACAAATCTAGGATTTGGCTCGGAACTTTCAAAACAG  
CTGAGATCGCAGCTCGTGCTCACGACGTCGCCGCTTAGCTCTCCGTGGAAGAGGCGCCT  
GCCTCAACTTCGCCGACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACCTGCGCCAAGG  
ATATCCAGAAGGCTGCTGCTGAAGCCGCAATTGGCTTTTGAGGCCGAGAAGAGTGATACCA  
CGACGAATGATCATGGCATGAACATGGCTTCTCAGGCTGAGGTTAATGACACGACGGATC  
ATGGCTCTGGACATGGAGGAGACGATGGTGGAGGCTGTTTTACTGAGGAGCAGAGAGACG  
GGTTTTACATGGCGGAGGAGACGCGTGGAGGCTGTTGTTCCGGAGGAACAGATGAGCA  
AAGGGTTTTACATGGACGAGGAGTGGATGTTCCGGATGCCGACCTTGTGGCTGATATGG  
CGGCAGGGATGCTCTTACCGCCGCGCTCCGTACAATGGGGACATAATGATGACTTCGAAG  
GAGATGCTGACATGAACCTCTGGAATTATTAGTACTCGTATTTTTTTAAATTATTTTTTG  
AACGAATAATATTTTTATTGAATTCGGATTCTACCTGTTTTTTTTAATGGAT

**b CBF5** Species=Brassica oleracea Length=877 [SEQ ID No. 72]  
ACCGCTCGAGCAACAATGAACACATTCCCTGCTTCCACTGAAATGGTTAGCTCCGAGAAC  
GAGTCTCCGGTTACTACGGTAGTAGGAGGTGATTATTATCCCATGTTGGCGGCAAGCTGT  
CCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGAGACACGTCACCCCATTTACCGAGGA  
GTTCTGCTGAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACAAACAAGAAATCT  
AGAATTTGGCTTGGAACTTTCAAAACAGCTGAGATGGCAGCTCGTGCTCACGACGTGGCT

SUBSTITUTE SHEET (RULE 26)

## FIGURE 18A-continued

GCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCTCAATTATGCGGACTCGGCTTGGCGGCTC  
CGCATCCCGGAGACAACCTGCCACAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATTG  
GCTTTTGGAGCTGAGAAGAGTGATGCGACGATGCAAAATGGCCTGAACATGGAGGAGACG  
ACGGCGGCGGCTTCTCAGACTGAAGTGAGTGACACGACGACAGATCATGGCATGAACATG  
GAGGAGACAACGGCGGTGGCTTCTCAGGCTGAGGTGAATGACACGACGACAGATCATGGC  
GTAGACATGGAGGAGACGATGGTGGAGGCTGTTTTTACTGAGGAACAAAGTGAAGGGTTC  
AACATGGCGAAGGAGTCGACGGCGGAGGCTGCTGTTGTTACGAGGAACTGAGCAAAGGA  
GTTTACATGGACGAGGAGTGGACGTACGAGATGCCGACCTTGTGGCTGATATGGCGGCA  
GGGATGCTTTTGGCGCCACCATCTGTACAATGGGGACATAATGATGACTTGGAAGGAGAT  
GCGGACATGAACCTACTGGAGTTATTAAGGATCCGCG

**brCBF1** Species=Brassica rapa Length=374 [SEQ ID No. 74]  
CATCCCATTACAGGGGGTTCGTTTAAAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGG  
GAACCAACAAGAAATCTAGGATTTGGCTCGGAACCTTTCAAACCGCTGAGATCGCTGCT  
CGTGCTCAGACGTTGCTGCCCTAGCCCTCCGCGGGAGAGGCGCCTGCCTCAACTTCGCC  
GACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACCTGCGCCAAGGACATCCAGAAGGCG  
GCTGCTGAAGCTGCATTGGCTTTTGGGCGGAGAAGAGTGATCATGGCATGAACATCAAG  
AATACTACGGCGGTGTTTCTCAGGTTGAGGTGAATGACACGACGACGACCACGGCTTG  
GACATGGAGGAGAC

**brCBF2** Species=Brassica rapa Length=884 [SEQ ID No. 76]  
TACACTCAGCCTTATCCAGTTTTTTTCAAAGACTTTTCAACAATGAACACATTCCCTGC  
GTCCACTGAAATGGTTGGCTCCGAGAACGAGTCTCCGGTTACTACGGTAGCAGGAGGTGA  
TTATTATCCCATGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCA  
GGAGACACGTCACCCCATTTACCGAGGAGTTTCGTCTGAGAAAGTCAGGTAAGTGGGTGTG  
TGAAGTGAGGGAACCAACAAGAAATCTAGAATTTGGCTTGGAACTTTCAAACACGCTGA  
GATGGCAGCTCGTGCTCAGCAGCTCGCTGCCCTAGCCCTCCGTGGAAGAGGCGCCTGCCT  
CAATTATGCGGACTCGGCTTGGCGGCTCCGCATCCCGGAGACAACCTGCCACAAGGATAT  
CCAGAAGGCTGCTGCTGAAGCCGCAATTGGCTTTTGGGCTGAGAAAAGTGATGTGACGAT  
GCAAAATGGCCTGAACATGGAGGAGATGACGGCGGTGGCTTCTCAGGCTGAAGTGAATGA  
CAGCAGACAGAACATGGCATGAACATGGAGGAGGCAACGGCAGTGGCTTCTCAGGCTGA  
GGTGAATGACACGACGACGATCATGGCGTAGACATGGAGGAGACAATGGTGGAGGCTGT  
TTTTACTGAGGAACAAGTGAAGGGTTAACATGGCGAAGGAGTCGACGGTGGAGGCTGC  
TGTTGTTACGGAGGAACCGAGCAAAGGATCTTACATGGACGAGGAGTGGATGCTCGAGAT  
GCCGACCTTGTTGGCTGATATGGCGGAAGGGATGCTTTTGGCGCGCCGTCCTGACAATG  
GGGACAGAATGATGACTTCCAAGGAGATGCTGACATGAACCTCT

**brCBF3** Species=Brassica rapa Length=806 [SEQ ID No. 78]  
ACACTCAGCCTTATCCAGTTTTTCAAAAAAGTATTCAACGATGAACCTCAGTCTCTACTTT  
TTCTGAACTGCTCTGCTCCGAGAACGAGTCTCCGGTTAATACGGAAGGTGGTGATTACAT  
TTTGGCGGCGAGCTGTCCCAAGAAACCTGCTGGTAGGAAGAAAGTTTCAGGAGACACGCCA  
CCCCATTTACAGAGGAGTTTCGTCTGAGGAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGA  
ACCAACAAGAAATCTAGAATTTGGCTCGGAACCTTTCAAACAGCTGAGATCGCAGCTCG  
TGCTCAGCAGCTTGCCGCTTAGCTCTCCGTGGAAGAGGCGCCTGCCTCAACTTCGCCGA  
CTCGGCTTGGCGGCTCCGTATCCCGGAGACGACCTGCGCCAAGGATATCCAGAAGGCTGC  
TGCTGAAGCCGCATTGGCTTTTGGGCGGAGAAGAGTGATACCACGACGAATGATCGTGG  
CATGAACATGGAGGAGACGTCGGCGGTGGCTTCTCCGGCTGAGTTGAATGATACGACGGC  
GGATCATGGCCTGGACATGGAGGAGACGATGGTGGAGGCTGTTTTAGGGAGGAACAGAG  
AGAAGGGTTTTACATGGCGGAGGAGACGACGGTGGAGGGTGTGTTCCGGAGTAACAGAT  
GAGCAAAGGGTTTTACATGGACGAGGAGTGGACGTTGAGATGCCGAGGTTGTTGGCTGA  
TATGGCGGAAGGGATGCTTTTGGCGCCCCGTCCGTACAATGGGGACATAACGATGACTT  
CGAAGGAGATGCTGACATGAACCTCT

**brCBF4** Species=Brassica rapa Length=755 [SEQ ID No. 80]  
ACCGCTCGAGTACTTACTATACTACACTCAGCCTTATCCAGTTTTTCTTCCAACGATGGA  
CTCAATCTCTACTTTTCTGAACTGCTTGGCTCAGAGAACGAGTCTCCGGTTACTACGGT  
AGTAGGAGGTGATTATTGTCCCAGGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCGGGTAG  
GAAGAAGTTTTCAGGAGACACGTCACCCCATTTACCGTGGAGTTGCTTTAAGAAAGTCCGG  
TAAGTGGGTGTGTGAAGTGAGGGAACCAACAAGAAATCTAGGATTTGGCTCGGAACCTTT  
CAAACCGCTGAGATCGCTGCTCGTGCTCAGCAGCTTGTGCTTAGCCCTCCGCGGAAG  
AGGCGCCTGCCTCAACTTCGCCGACTCGGCTTGACGGCTCCGTATCCCGGAGACAACCTG  
CGCCAAGGATATCCAGAAGGCTGCTGCTGAAGCTGCATTGGCTTTTGGAGCCGAGAAGAG  
TGATCATGGCATGAACATGAAGAATACTACGGCGGTGGCTTCTCAGGTTGAGGTGAATGA  
TACGACGACGACCATGGCGTGGACATGGAGGAGACGAGGTTGGAGGTTGTTGTACGGA  
GGAACAGAACAATGGTTTTACATGGACGAGGAGTGGATGTTTGGGATGCCGACGTTGTT  
GTTGATATGGCGGAAGGGATGCTTATACCGCGCAGTCCGTACAATCGGGACACTACGA

SUBSTITUTE SHEET (RULE 26)

**FIGURE 18A-continued**

TGACTTCGAAGGAGATGCTGACATGAACCTCTGGA

**brCBF5** Species=Brassica rapa Length=832 [SEQ ID No. 82]  
ACCGCTCGAGTACTTACTATACTACACTCAGCCTTATCCAGTTTTTCTTCCAACGATGGA  
CTCAATCTCTACTTTTCTGAACTGCTTGGCTCAGAGAACGAGTCTCCGGTTACTACGGT  
AGTAGGAGGTGATTATTGTCCCAGGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCGGGTAG  
GAAGAAGTTTTCAGGAGACACGTACCCCATTTACCGTGGAGTTCGTTTAAGAAAGTCCGG  
TAAGTGGGTGTGTGAAGTGAGGGAACCAACAAGAAATCTAGGATTTGGCTCGGAACTTT  
CAAAACCGCTGAGATCGCTGCTCGTGTCTCAGACGTTGCTGCCTTAGCCCTCCGCGGAAG  
AGGCGCTGCCTCAACTTCGCCGACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACCTG  
CGCCAAGGATATCCAGAAGGCTGCTGCTGAAGCTGCTTTGGCTTTTGAGGCCGAGAAGAG  
TGATCATGGCATGAACATGAAGAATACTACGGCGGTGGCTTCTCAGGTTGAGGTGAATGA  
TACGACGACGACCATGGCGTGGACATGGAGGAGACGTTGGTGGAGGCTGTTTTTACGGA  
GGAACAGAGAGAAGGGTTTTACATGACGGAGGAGACGAGGGTGGAGGGTGTGTTTACGGA  
GGAACAGAACAAATTGGTTTTACATGGACGAGGAGTGGATGTTTGGGATGCCGACGTTGTT  
GGTTGATATGGCGGAAGGGATGCTTATACCGCGGCAGTCCGTACAATCGGGACACTACGA  
TGACTTCGAAGGAGATGCTGACATGAACCTCTGGAATTATTAGGGATCCGCG

**brCBF6** Species=Brassica rapa Length=830 [SEQ ID No. 84]  
TACTACACTCAGCCTTATCCAGTTTTCAAAAAAAGTATTCAACTATGAACCTCAGTCTCTA  
CTTTTTCTGAACTGCTCTGCTCCGAGAACAAGTCTCCGGTTAATACGGAAGGTGGTGATT  
ACATTTTGGCGGCGAGCTGTCCCAAGAAACCTGCTGGTAGGAAGAAGTTTACAGGAGACAC  
GCCACCCCATTTACAGAGGAGTTCGCCTAAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGA  
GGGAACCAACAAGAAATCTAGAATTGGCTCGGAACTTCAAAACAGCTGAGATAGCAG  
CTCGTGCTCAGACGTCGCCGCTTAGCTCTCCGTGGAAGAGGCGCTGCCTCAACTTCG  
CCGACTCGGCTTGGCGGCTCCGTATCCAGAGACAACCTGCGCCAAGGATATCCAGAAGG  
CTGCTGCTGAAGCCGATTGGCTTTTGAGGCCGAGAAGAGTGATACCACGACGAATGATC  
GTGGCATGAACATGGAGGAGACGTCGCCGCTGGCTTCTCCGCTGAGTTGAATGATACGA  
CGGCGGATCATGGCCTGGACATGGAGGAGACGATGGTGGAGGCTGTTTTAGGGACGAAC  
AGAGAGAAGGGTTTTACATGGCGGAGGAGACGACGGTGGAGGGTGTGTTCCGGAGGAAC  
AGATGAGCAAAGGGTTTTACATGGACGAGGAGTGGACGTTCCGAGATGCCGAGGTTGTTGG  
CTGATATGGCGGAAGGGATGCTTCTGCCTCCCCCGTCCGTACAATGGGGACATAACGATG  
ACTTCGAAGGAGATGCTGACATGAACCTCTGGAATTATTAGGGATCCGCG

**brCBF7** Species=Brassica rapa Length=854 [SEQ ID No. 86]  
CTATACTACACACAGCCTTATCCAGCCGCTCGAGTACTTACTATACTACACTCAGCCTTT  
TCCAGTTTTTCAAAAGAAGTTTTCAACGATGAATCAGTCTCTACTCTTTCTGAAGTTCT  
TGGCTCCCAGAACGAGTCTCCCGTAGGTGGTGATTACTGTCCCATGTTGGCGGCGAGCTG  
TCCGAAGAAGCCGGCGGGTAGGAAGAAGTTTCGGGAGACACGTACCCCATTTACAGAGG  
AGTTCGTCTTAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGAGGGAACCAACAAGAAATC  
TAGGATTTGGCTCGGAACTTTCAAAACAGCTGAGATCGCAGCTCGTGCTCAGACGTTGC  
CGCCTTAGCTCTCCGTGGAAGAGGCGCTGCCTCAACTTCGCCGACTCGGCTTGGCGGCT  
CCGTATCCCGGAGACAACCTGCGCCAAGGATATCCAGAAGGCTGCTGCTGAAGCCGCATT  
GGCTTTTGAGGCGGAGAAGAGTGATACCACGACGACGAATGATCATGGCATGAACATGGC  
TTCTCAGGTTGAGGTTAATGACACGACGGATCATGACCTGGACATGGAGGAGACGATGGT  
GGAGGCTGTTTTAGGGAGGAACAGAGAGAAGGGTTTTACATGGCGGAGGAGACGACGGT  
GGAGGGTATTGTTCCGGAGGAACAGATGAGCAAAGGGTTTTACATGGACGAGGAGTGGAT  
GTTCCGGATGCCGACCTTGTGGCTGATATGGCGGCAGGGATGCTCTTACCGCCGCCGCTC  
CGTACAATGGGGACATAATGATGACTTCGAAGGAGATGCTGACATGAACCTCTGGAATTA  
TTAAGGGATCCGCG

**gmCBF1** Species=Glycine max Length=738 [SEQ ID No. 88]  
CATCCGATTTATAGTGGCGTGAGGAGGAGGAACACGGATAAGTGGGTAAGTGAGGTGAGG  
GAGCCCAACAAAAAGACCAGGATTTGGCTGGGGACTTTTCCACGCCGGAGATGGCGGCA  
CGGGCCACGACGTGGCGGCAATGGCCCTGAGGGGCCGGTATGCCTGTCTCAACTTCGCT  
GACTCGACGTGGCGGTTACCAATTCCCGCCACTGCTAACGCAAAGGATATACAGAAAGCA  
GCAGCAGAGGCTGCCGAGGCTTTCAGACCAAGTCAGACCTTAGAAAATACGAATACAAAG  
CAAGAGTGTGTAAGAGCAAGTGTAGATATGCCTGAGTTGCTTAGGAATATGGTGCTT  
TATACGGAGGAAGAAAGTGTGACGACAACAACGATCACAGAACAAAAACGAGGAATGTTT  
ATGTCCCCAACACATTGCATAGGGTATGAGTATGAAGATGCTGACTTGGATGCTCAAGAT  
GCTGAGGTGTCCCTATGGAGTTTCTCAATTTAATAACGTGCTTTTGGTTTGGTTTTTAT  
GTTAGTTTTGGAGTGTGACTGTCTGTACTGGTTTTTTATTAGTAGTACGGATACTAGCTA  
TAGGTGGCAGATTGAAAGGGACCAAAAGGAATTTCTTTTGAAACCCTTTTTGTCAAAGT  
AATCAATCGCGTATCATCAAGTGAATCCCTTGATCAAGTTTATGTATGAATTAATAAAA  
GAAGAATCTAGTTTTGGT

**SUBSTITUTE SHEET (RULE 26)**

**FIGURE 18A-continued**

**rsCBF1** Species= *Raphanus sativus* Length=793 [SEQ ID No. 90]  
ACTACACTCAGCCTTATCCAGTTTTCTTCCAACGATGGACTCAATCTCTACTTTTTCTG  
AACTGCTTGGCTCCGAGAACGAGTCTCCGGTTACTACGGTAGTAGGAGGTGATTATTTTC  
CCAGGTTGGCGGCAAGCTGTCCGAAGAAGCCAGCGGGTAGGAAGAAGTTTCAGGAGACAC  
GTCACCCCATTTACCGCGGAGTTCTGTTTAAAGAAAGTCAGGTAAGTGGGTGTGTGAAGTGA  
GGGAACCAAACAAGAAATCTAGGATTTGGCTCGGAACTTTCAAAACCGCTGAGATCGCTG  
CTCGTGCTCACGACGTTGCTGCCTTAGCCCTCCGCGGAAGAGGCGCCTGCCTCAACTTCG  
CCGACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACCTGCGCCAAGGATATCCAGAAGG  
CTGCTGCTGAAGCTGCATTGGCTTTTGAGGCCGAGAAGAGTGATCATGGCATGAACATGA  
AGAATACTACGGCGGTGGCTTCTCAGGTTGAGGTGAATGACACGACGACGGACCATGGCG  
TGGACATGGAGGAGACGTTGGTGGAGGCTGTTTTTACGGAGGAACAGAGAGAAGGGTTTT  
ACATGACGGAGGAGACGAGGGTGGAGGGTGTGTGTACGGAGGAACAGAACAATTGGTTTT  
ACATGGACGAGGAGTGGATGTTTGGGATGCCGACGTTGTTGGTTGATATGGCGGAAGGGA  
TGCTTTTACCGCGGCCGTCCGTACAATCGGGACACTACGATGACTTCGAAGGAGATGCTG  
ACATGAACCTCTG

**rsCBF2** Species= *Raphanus sativus* Length=682 [SEQ ID No. 92]  
ACACCTAAACCTTATCCAGGTTTAACTTTTTTTTTCATAAAGAGTTTTCAACAATGACCA  
CATTTTCTACCTTTTCCGAAATGTTGGGCTCCGAGTACGAGTCTCCGGTTACATTAGGCG  
GAGAGTATTGTCCGAAGCTGGCCGCGAGCTGTCCGAAGAAACCAGCTGGTCGTAAGAAGT  
TTCGAGAGACGCGCCACCCAATATACAGAGGAGTTCTGCTGAGAAACTCAGGTAAGTGGG  
TGTGTGAAGTGAGGGAGCCAAACAAGAAATCTAGGATTTGGCTCGGTACTTTCTTAACCG  
CCGAGATCGCAGCGCGTGCCACGACGTCGCCGCCATAGCCCTCCGCGGCAAAATCCGCAT  
GTCTCAATTTTCGCTGACTCGGCTTGGCGGCTCCGTATCCCGGAGACAACATGCCCCAAGG  
ATATACAGAAGGCGGCTGCTGAAGCCGCGGTGGCTTTTTCAGGCTGAGATAAATGATACGA  
CGACGGATCATGGCCTGGACTTGGAGGAGACGATCGTGGAGGCTATTTTTACGGAGGTAA  
ACAACGATGAGTTTTATATGGACGAGGAGTCCATGTTTCGGGATGCCGTCTTTGTTGGCTA  
GTATGGCGGAAGGGATGCTTTTGCCGCTGCCGTCCGTACAATCTGAACATAACTGTGACT  
TCGACGGAGATGCTGACATGAA

**zmCBF1** Species= *Zea mays* Length=349 [SEQ ID No. 94]  
CGGAGTCCGCGGACGGCGGGCGGGCGGGCGGACGACGAGTACGCGACGGTGCTGTCGGCGC  
CACCCAAGCGGCGGGCGGGCGGGACCAAGTTCCGGGAGACGCGGCACCCCGTGTAACGCG  
GCGTGCGGCGGGCGGGCGGGCGGGCGGGCGGTGTGCGAGGTCCGCGAGCCCAACAAGA  
AGTCGCGCATCTGGCTCGGCACCTTCGCCACCCCGAGGCCGCGCGCGCGCACGACG  
TGGCCGCGCTGGCCCTGCGGGGCGCGCGCGGTGCCTCAACTTCGCCGACTCGGCGCGCC  
TGCTCCAAGTCGACCCCGCCACGCTCGCCACCCCGACGACATCCGCCG

**FIGURE 18B**

**BJCBF1-PEP** Species=Brassica juncea length=130 [SEQ ID No. 39]  
 LPGVRLRKSGKWVCEVREPNNKSRIWLGTFKTAIEAARAHDAALALRGRAACLNFDASA  
 WRLRIPETTCAKDIQKAAAEAAALAFGAEKSDTTTNDQGMNMEEMTAVASQAEVSDTTTYH  
 GLDMEETMVD

**BJCBF2-PEP** Species=Brassica juncea length=143 [SEQ ID No. 41]  
 HPIYRGVRLRKSGKWVCEVREPNNKSRIWLGTFKTAIEAARAHDAALALRGKSACLNFA  
 DSAWRLRIPETTCPKIEIQKAAAEAAVAFQAE LNDDTTADHGLDVEETIVEAIFTEESSEGF  
 YMDEEFMFGMPTLWASMAEGMLL

**BJCBF3-PEP** Species=Brassica juncea length=143 [SEQ ID No. 43]  
 HPIYRGVRLRKSGKWVCEVREPNNKSRIWPGTFLTAIEAARAHDAALALRGKSACLNFA  
 DSAWRLRIPETTCPKIEIQKAAAEAAVAFQAE LNDDTTADHGLDVEETIVEAIFTEESSEGF  
 YMAEEFMFGMPTLWASVAEGMLL

**BJCBF4-PEP** Species=Brassica juncea length=142 [SEQ ID No. 45]  
 HPIYRGVRLRKSGKWVCEVREPNNKSRIWLGTFPTVEMAARAHDAALALRGKSACLNFA  
 DSAWRLRIPETTCPKIEIQKAAAEAAVAFQAE LNDDTTADHGLDVEETIVEAIFTEESSEGF  
 ENGVEYMDPRFLENMAEGMLF

**BNCBF1-PEP** Species=Brassica napus length=210 [SEQ ID No. 47]  
 HPIYRGVRLRKSGKWVCEVREPNNKSRIWLGTFKTAEMAARAHDAALALRGKGACLNFA  
 DSAWRLRIPETTCBKDIQKAAAEAAALAFEAESDVTMONGNMEETTAVASQAEVNDTTT  
 EHGMMNEEATAVASQAEVNDTTTTHGVDMEETMVEAVFTGEQSEGFMMAKESTVEAAVVT  
 EEPKSGSYMDEEWMLEMTLLADMAEGMLL

**BNCBF2-PEP** Species=Brassica napus length=283 [SEQ ID No. 49]  
 MNTFPASTEMVGSSENPVTTTVGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
 SGKWVCEVREPNNKSRIWPGTFTKTAEMAARAHDAALALRGKGACLNFAADSAWRLRIPET  
 TCHKDIQKAAAEAAALAFEAESDVTMONGNMEETTAVASQAEVNDTTTTHGMMNEEATA  
 VASQAEVNDTTTTHGVDMEETMVEAVFTGEQSEGFMMAESTVEAAVVTDELKSGFYMD  
 EWTYEMPTLLADMAAGMLLPPPSVQWGHNDDEGDADMNLWSY

**BNCBF3-PEP** Species=Brassica napus length=279 [SEQ ID No. 51]  
 MNTFPASTEMVGSSENPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
 SGKWVCEVREPNNKSRIWPGTFTKTAEMAARAHDAALALRGKGACLNFAADSAWRLRIPET  
 TCHKDIQKAAAEAAALAFEAESDVTMONGNMEETTAVASQAEVNDTTTTHGMMNEEATA  
 VASQAEVNDTTTTHGVDMEETMVEAVFTGEQSEGFMMAKESTVEAAVVTDEPSKGSYMDE  
 EWMLEMTLLADMAEGMLLPPPSVQWGHNDDEGDADMN

**BNCBF4-PEP** Species=Brassica napus length=250 [SEQ ID No. 53]  
 MNSVSTFSELLGSENPVGGDYCPMLAASCPKPKPAGRKKFRETRHPIYRGVRLRKSGKW  
 VCEVREPNNKSRIWLGTFKTAIEAARAHDAALALRGKGACLNFAADSAWRLRIPETTCAK  
 DIQKAAAEAAALAFEAESDTTTNDHGMNMAEQAEVNDTTDHDLDMEETMVEAVFTTEEQD  
 GFYMAEETTVEGVVPEEQMSKGFYMDDEEWMFGMPTLLADMAAGMLLPPPSVQWGHNDDE  
 GDVDMNLWNY

**BNCBF5-PEP** Species=Brassica napus length=251 [SEQ ID No. 55]  
 MNSVSTFSELLRSENPVNTTEGGDYILAASCPKPKPAGRKKFQETRHPIYRGVRLRKSGK  
 WVCEVREPNNKSRIWLGTFKTAIEAARAHDAALALRGKGACLNFAADSAWRLRIPETTC  
 KDIQKAAAEAAALAFEAESDTTTNDHGMNMAEQAEVNDTTDHDLDMEETIVEAVFREEQ  
 EGFYMAEETTVEGVVPEEQMSKGFYMDDEEWMFGMPTLLADMAAGMLLPLPSVQWGHNDDE  
 EGDADMNLWNY

**BNCBF6-PEP** Species=Brassica napus length=277 [SEQ ID No. 57]  
 MNTFPASTEMVGSSENPVTTTVGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
 SGKWVCEVREPNNKSRIWLGTFKTAEMAARAHDAALALRGKGACLNFAADSAWRLRIPET  
 TCHKDIQKAAAEAAALAFEAESDVTMEETMAVASQAEVNDTTTTHGMMNEEATAVASQAE  
 VNDTTTTHGVDMEETMVEAVFTTEEQSEGFMMAESTVEAAVVTDELKSGFYMDDEEWTYEM  
 PTLLADMAAGMLLPPPSVQWGHNDDEGDADMNLWNY

**BNCBF7-PEP** Species=Brassica napus length=213 [SEQ ID No. 59]  
 MNSVSTFSELLGSENPVGGDYCPMLAASCPKPKPAGRKKFRETRHPIYRGVRLRKSGKW  
 VCEVREPNNKSRIWLGTFKTAIEAARAHDAALALRGKSACLNFAADSAWRLRIPETTCPK  
 IEQKAAAEAAVAFKAEINNTTADHGIDVEETIVEAIFTEENNDGFYMDDEESMFGMPALL

**SUBSTITUTE SHEET (RULE 26)**



## FIGURE 18B-continued

ASMAEGMLLPPPSVQFGHTYDFDGDADVSLWSY

**BNCBF8-PEP** Species=Brassica napus length=277 [SEQ ID No. 61]  
MNTFPASTEMVGSENE SPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAEMAARAHDAALALRGRGACLNADYASASRLRIPET  
TCHKDIQKAAAEALAFEAEKSDVTMEETMAVASQAEVNDTTDGHMNMEEATAVASQAE  
VNDTTTDHGVDMETMVEAVFTGEQSEGFNMAKESTVEAAVVTEEPSKGSYMDDEEWMLM  
PTLLADMAEGMLLPPPSVQWQNDDEFGDADMNLWSY

**BNCBF9-PEP** Species=Brassica napus length=283 [SEQ ID No. 63]  
MNTFPASTEMVGSENE SPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAEMAARAHDAALALRGRGARLNADYASAWRLRIPET  
TCHKDIQKAAAEALAFEAEKSDVTMONGLNMEETAVASQAEVNDTTTEHGMNMEEATA  
VASQAEVNDTTTDHGVDMETMVEAVFTGEQSEGFNMAKESTVEAAVVTEEPSKGSYMD  
EWMLEMPPTLLADMAEGMLLPPPSVQWQNDDEFGDAHMNLWSY

**BOCBF1-PEP** Species=Brassica olercea Length=188 [SEQ ID No. 65]  
HPIYRGVRLRKSGKWVCEVREPNNKKSRIWLGTFKTAEMIAARAHDAALALRGRACLNFA  
DSAWRLRIPETTCAKDIQKAAAEALAFGAEKSDTTNDQGMNMEEMTVVASQAEVSDTT  
TYHGLDMEETMVEAVFAEEQREGFYLAEEETTVEGVVTEEQSKGFYMDDEWTFGMQSFLAD  
MAEGMLFP

**BOCBF2-PEP** Species=Brassica olercea Length=152 [SEQ ID No. 67]  
MTSFSTFSELLGSEHESPVTLGEEYCPKLAASCPKPKPAGRKKFRETRHPVYRGVRLRNSG  
KWVCEVREPNNKKSRIWLGTFKTAEMIAARAHDAALALRGRGACLNFAADSAWRLRIPETTC  
PKEIQKAAAEAAVAFKAEINNTTADHGLDMEE

**BOCBF3-PEP** Species=Brassica olercea Length=277 [SEQ ID No. 69]  
MNTFPASTEMVGSENE SPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAEMAARAHDAALALRGRGACLNADYASAWRLRIPET  
TCHKDIQKAAAEALAFEAEKSDVTMEETMAVASQAEVNDTTDGHMNMEEATAVASQAE  
VNDTTTDHGVDMETMVEAVFTEEQSEGFNMAEESTVEAAVVTEELSKGFYMDDEEWTYEM  
PTLLADMAAGMLLPPPSVQWGHNDDEFGDADMNLWSY

**BOCBF4-PEP** Species=Brassica olercea Length=250 [SEQ ID No. 71]  
MNSVSTFSELLGSENE SPVGGDYCPMLAASCPKPKPAGRKKFRETRHPIYRGVRLRKSGKW  
VCEVREPNNKKSRIWLGTFKTAEMIAARAHDAALALRGRGACLNFAADSAWRLRIPETTC  
DIQKAAAEALAFEAEKSDTTNDHGMNMASQAEVNDTTDGHGLDMEETMVEAVFTEEQORD  
GFYMAEETTVEGVVPEEQMSKGFYMDDEEWMFGMPTLLADMAAGMLLPPPSVQWGHNDDE  
GDADMNLWN

**BOCBF5-PEP** Species=Brassica olercea Length=287 [SEQ ID No. 73]  
MNTFPASTEMVSSSENE SPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVRELNNKKSRIWLGTFKTAEMAARAHDAALALRGRGACLNADYASAWRLRIPET  
TCHKDIQKAAAEALAFEAEKSDATMONGLNMEETAAASQTEVSDTTTDHGMNMEEATA  
VASQAEVNDTTTDHGVDMETMVEAVFTEEQSEGFNMAKESTAEAAVVTEELSKGFYMD  
EWTYEMPTLLADMAAGMLLPPPSVQWGHNDDEFGDADMNLLELLRIR

**BRCBF1-PEP** Species=Brassica rapa Length=124 [SEQ ID No. 75]  
HPIYRGVRLRKSGKWVCEVREPNNKKSRIWLGTFKTAEMIAARAHDAALALRGRGACLNFA  
DSAWRLRIPETTCAKDIQKAAAEALAFEAEKSDHGMNKNITTA VVSQAEVNDTTTDHGL  
DMEE

**BRCBF2-PEP** Species=Brassica rapa Length=280 [SEQ ID No. 77]  
MNTFPASTEMVGSENE SPVTTVAGGDYYPMLAASCPKPKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAEMAARAHDAALALRGRGACLNADYASAWRLRIPET  
TCHKDIQKAAAEALAFEAEKSDVTMONGLNMEETAVASQAEVNDTTTEHGMNMEEATA  
VASQAEVNDTTTDHGVDMETMVEAVFTEEQSEGFNMAKESTVEAAVVTEEPSKGSYMD  
EWMLEMPPTLLADMAEGMLLPPPSVQWQNDDEFGDADMNL

**BRCBF3-PEP** Species=Brassica rapa Length=204 [SEQ ID No. 79]  
MNSVSTFSELLCSENE SPVNTTEGGDYILAASCPKPKPAGRKKFQETRHPIYRGVRLRKSGK  
WVCEVREPNNKKSRIWLGTFKTAEMIAARAHDAALALRGRGACLNFAADSAWRLRIPETTC  
KDIQKAAAEALAFEAEKSDTTNDRGMNMEETSAVASPAELNDTTADHGLDMEETMVEA  
VFREEQREGFYMAEETTVEGVVPE

SUBSTITUTE SHEET (RULE 26)

**FIGURE 18B-continued**

**BRCBF4-PEP** Species=Brassica rapa Length=112 [SEQ ID No. 81]  
MDSISTFPELLGSENEPVTTVVGGDYCPRLAASCPKKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSA

**BRCBF5-PEP** Species=Brassica rapa Length=255 [SEQ ID No. 83]  
MDSISTFPELLGSENEPVTTVVGGDYCPRLAASCPKKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSAWRLRIPET  
TCAKDIQKAAAEAAALAFEAEKSDHGMNMKNNTTAVASQVEVNDTTTDHGVDMEEETLVEAVF  
TEEQREGFYMTTEETRVGVVTEEQNNWFYMDDEEWMFGMPTLLVDMAGMLIPRQSVQSGH  
YDDFEGDADMNLWNY

**BRCBF6-PEP** Species=Brassica rapa Length=258 [SEQ ID No. 85]  
MNSVSTFSELLCSENKSPVNTTEGGDYILAASCPKKPAGRKKFQETRHPIYRGVRLRKSGK  
WVCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSAWRLRIPETTC  
KDIQKAAAEAAALAFEAEKSDTTTNDRGMNMEETSAVASPAELNDTTADHGLDMEETMVEA  
VFRDEQREGFYMAEETTVEGVVPEEQMSKGFYMDDEEWTFFEMPRLLADMAEGMLLPPPSVQ  
WGHNDDEFGDADMNLWNY

**BRCBF7-PEP** Species=Brassica rapa Length=251 [SEQ ID No. 87]  
MNSVSTLSEVLGSEQNESPVGGDYCPMLAASCPKKPAGRKKFQETRHPIYRGVRLRKSGKW  
VCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSAWRLRIPETTC  
KDIQKAAAEAAALAFEAEKSDTTTNDHGMNMAEQVEVNDTTDHDLDMEETMVEAVFREEQ  
EGFYMAEETTVEGIVPEEQMSKGFYMDDEEWMFGMPTLLADMAAGMLLPPPSVQWGHNDDE  
EGDADMNLWNY

**GMCBF1-PEP** Species=Glycine max Length=170 [SEQ ID No. 89]  
HPIYSGVRRRNTDKWVSEVREPNNKTRIWLGTFTPEMAARAHDVAAMALRGRYACLNFA  
DSTWRLPIPATANAKDIQKAAAEAAEAFFRPSQTLENTNTKQECVKVTTTTITEQKRGMF  
YTEEEEQVLDMPPELLRNMLVMSPTHICIGYEYEDADLDAQDAEVSLSWFSI

**RSCBF1-PEP** Species=Raphanus sativus Length=252 [SEQ ID No. 91]  
MDSISTFSELLGSENEPVTTVVGGDYFPRLAASCPKKPAGRKKFQETRHPIYRGVRLRK  
SGKWVCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSAWRLRIPET  
TCAKDIQKAAAEAAALAFEAEKSDHGMNMKNNTTAVASQVEVNDTTTDHGVDMEEETLVEAVF  
TEEQREGFYMTTEETRVGVVTEEQNNWFYMDDEEWMFGMPTLLVDMAGMLLPRPSVQSGH  
YDDFEGDADMNL

**RSCBF2-PEP** Species=Raphanus sativus Length=209 [SEQ ID No. 93]  
MTTFSTFSEMLGSEYESPVTLGGEYCPKLAASCPKKPAGRKKFQETRHPIYRGVRLRNSG  
KWVCEVREPNNKKSRIWLGTFKTAETIAARAHDVAALALRGRGACLNFAADSAWRLRIPETTC  
PKDIQKAAAEAAAFQAEINDTTTDHGLDLEETIVEAIFTEVNDEFYMDDEESMFGMPSL  
LASMAEGMLLPLPSVQSEHNCDFDGDADM

**ZMCBF1-PEP** Species=Zea maize Length=115 [SEQ ID No. 95]  
ESADGGGGGDDEYATVLSAPPKRPAGRTKFRRETRHPVYRGVRRRGPAWRVCEVREPNNK  
SRIWLGTFATPEAARAHDVAALALRGRAACLNFAADSAARLLQVDPATLATPDDIR

FIGURE 19A

ap2{atcbf2}	1	HpiYrGVRqR	.nsgkwVcEi	RepNkK.tRI	WlGTFqTaEm	AArAhDVAAi	50	ALRGsAcLN	64	fADS
ap2{atcbf3}		HpiYrGVRrR	.nsgkwVcEv	RepNkK.tRI	WlGTFqTaEm	AArAhDVAAi		ALRGsAcLN		fADS
ap2{atcbf1}		HpiYrGVRqR	.nsgkwVsEv	RepNkK.tRI	WlGTFqTaEm	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bjcbf4}		HpiYrGVRqR	.nsgkwVcEv	RepNkK.sRI	WlGTFpTvEm	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bocbf2}		HPvYrGVRlR	.nsgkwVcEv	RepNkK.sRI	WlGTFlTaEi	AArAhDVAAi		ALRGsAcLN		fADS
ap2{rscbf2}		HpiYrGVRlR	.nsgkwVcEv	RepNkK.sRI	WlGTFlTaEi	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bjcbf2}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFlTaEi	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bncbf7}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFlTaEi	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bjcbf3}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WpGTFlTaEi	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bncbf2}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WpGTFkTaEm	AArAhDVAAi		ALRGsAcLN		fADS
ap2{bncbf3}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WpGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf1}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf6}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf8}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf3}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf2}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf4}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf5}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bocbf4}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf1}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf3}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf4}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf5}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf6}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{brcbf7}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{rscbf1}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bocbf1}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEi	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bocbf5}		HpiYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{bncbf9}		HPvYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WpGTFkTaEm	AArAhDVAAi		ALRGsAcLN		YADS
ap2{zncbf1}		HPvYrGVRlR	.ksgkwVcEv	RepNkK.sRI	WlGTFkTaEa	AArAhDVAAi		ALRGsAcLN		YADS
ap2{gmcbf1}		HpiYsGVR.R	.rntdkWVsEv	RepNkK.tRI	WlGTFpTpEm	AArAhDVAAi		ALRGsAcLN		YADS
Consensus		HP-Y-GVR-R	-----WV-E-RE-NK----	RI	W-GTF-T-E-	AArAhDVAAi		ALRG--A-LN		-ADS
ap2{erebp2}		grhYrGVRqR	p.wgkfaaEi	RdpakngaRv	WlGTYeTaEe	AAIaYDKAAy		rmRGskALIN		fphr

FIGURE 19B

1	ap2{atcbf2}	HPIYrGVRQR	.nsgkwVcE1	REpNKK.tRI	WlGTFqTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	64
	ap2{atcbf3}	HPIYrGVRrR	.nsgkwVcEv	REpNKK.tRI	WlGTFqTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	50
	ap2{atcbf1}	HPIYrGVRQR	.nsgkwVsEv	REpNKK.tRI	WlGTFqTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bjcbf4}	HPIYrGVRQR	.nsgkwVcEv	REpNKK.sRI	WlGTFpTvEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bocbf2}	HPvYrGVRlR	.nsgkwVcEv	REpNKK.sRI	WlGTFtTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{rscbf2}	HPIYrGVRlR	.nsgkwVcEv	REpNKK.sRI	WlGTFtTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bjcbf2}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFtTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf7}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFtTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bjcbf3}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WpGTFtTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf2}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WpGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf3}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WpGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf1}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf6}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf8}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bocbf3}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf2}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf4}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf5}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bocbf4}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf1}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf3}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf4}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf5}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf6}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{brcbf7}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{rscbf1}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bocbf1}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bocbf5}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEi	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{bncbf9}	HPIYrGVRlR	.ksgkwVcEv	REpNKK.sRI	WlGTFkTaEm	AARAHDVAAI	ALRGrsAcLN	fADS	
	ap2{zmcbf1}	HPvYrGVRrR	gpagrWVcEv	REpNKK.sRI	WlGTFaTpEa	AARAHDVAAI	ALRGraAcLN	fADS	
	ap2{gmcbf1}	HPIYSGVR.R	rintdkWVsEv	REpNKK.tRI	WlGTFpTpEm	AARAHDVAAI	ALRGryAcLN	fADS	
	ap2{dreb2a}	rcsfrGVRQR	i.wgkWaEi	REpnrG.sRI	WlGTFpTaqe	AASyDeAAK	AmyGplArLN	fprs	
	ap2{dreb2b}	hcsfrGVRQR	i.wgkWaEi	REpkig.tRI	WlGTFpTaek	AASyDeAAK	AmyGslArLN	fpqs	
	Consensus	-----GVR-R	-----WV-E-RE	-----RI	W-GTF-T---	AA-A-D-AA-	A--G--A-LN	---S	

FIGURE 19C

1	ap2 {atcbf2}	HPiYrGVRqR	.nsgkwwcEi	REpNkk.tRI	WlGTFqTaEm	AARAHdVAAi	ALRGrsAcLN	fADS	64
	ap2 {atcbf3}	HPiYrGVRrR	.nsgkwwcEi	REpNkk.tRI	WlGTFqTaEm	AARAHdVAAi	ALRGrsAcLN	fADS	
	ap2 {atcbf1}	HPiYrGVRqR	.nsgkwwsEv	REpNkk.tRI	WlGTFqTaEm	AARAHdVAAi	ALRGrsAcLN	fADS	
	ap2 {bjcbf4}	HPiYrGVRqR	.nsgkwwcEv	REpNkk.sRI	WlGTFpTvEm	AARAHdVAAi	ALRGrsAcLN	fADS	
	ap2 {bocbf2}	HPvYrGVRlR	.nsgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {rscbf2}	HPiYrGVRlR	.nsgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {bjcbf2}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {bncbf7}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {bncbf3}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {bncbf2}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WpGTFfLTaEi	AARAHdVAAi	ALRGksAcLN	fADS	
	ap2 {bncbf3}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WpGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bncbf1}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WpGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bncbf6}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bncbf8}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bocbf3}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {brcbf2}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bncbf4}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {bncbf5}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {bocbf4}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf1}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf3}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf4}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf5}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf6}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {brcbf7}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {rscbf1}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {bocbf1}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {bocbf5}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	fADS	
	ap2 {bncbf9}	HPiYrGVRlR	.ksgkwwcEv	REpNkk.sRI	WlGTFfLTaEi	AARAHdVAAi	ALRGgAcLN	yADS	
	ap2 {zmcbf1}	HPvYrGVRrR	gpagrWVcEv	REpNkk.sRI	WlGTFaTpEa	AARAHdVAAi	ALRGraAcLN	fADS	
	ap2 {gmcbf1}	HPiYsGVR.R	rntdkWVsEv	REpNkk.tRI	WlGTFpTpEm	AARAHdVAAi	ALRGyAcLN	fADS	
	ap2 {dreb2a}	rcsfrGVRqR	i.wgkwwaEi	REpnrG.sRI	WlGTFpTaqe	AAsAyDeAAk	amyGplArLN	fprs	
	ap2 {dreb2b}	hcsfrGVRqR	i.wgkwwaEi	REpkiG.tRI	WlGTFpTaeK	AAsAyDeAAk	amyGplArLN	fpqs	
	ap2 {tiny}	hpvYrGVRkR	.nwgkwwsEi	REpRkK.sRI	WlGTFpspeM	AARAhDVAAi	sikGasAiLN	fpDI	
	Consensus	-----GVR-R	-----WV-B	RE-----RI	W-GTF-----	AA-A-D-AA	---G--A-LN	----	

**FIGURE 19D**

1				50	64	
ap2{atcbf2}	HpiYrGVRqR .nsgkwwcEv	REpNkK.tRI	WlGTFqTaEm	AARAHdVAAI	ALGRsAcLN	FADs
ap2{atcbf3}	HpiYrGVRrR .nsgkwwcEv	REpNkK.tRI	WlGTFqTaEm	AARAHdVAAI	ALGRsAcLN	FADs
ap2{atcbf1}	HpiYrGVRqR .nsgkwwsEv	REpNkK.tRI	WlGTFqTaEm	AARAHdVAAI	ALGRsAcLN	FADs
ap2{bjcbf4}	HpiYrGVRqR .nsgkwwcEv	REpNkK.sRI	WlGTFpTvEm	AARAHdVAAI	ALGRsAcLN	FADs
ap2{bocbf2}	HPvYrGVRlR .nsgkwwcEv	REpNkK.sRI	WlGTFtTaEi	AARAHdVAAI	ALGRksAcLN	FADs
ap2{rscbf2}	HpiYrGVRlR .nsgkwwcEv	REpNkK.sRI	WlGTFtTaEi	AARAHdVAAI	ALGRksAcLN	FADs
ap2{bjcbf2}	HpiYrGVRlR .ksgkwwcEv	REpNkR.sRI	WlGTFtTaEi	AARAHdVAAI	ALGRksAcLN	FADs
ap2{bncbf7}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFtTaEi	AARAHdVAAI	ALGRksAcLN	FADs
ap2{bjcbf3}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WpGTfLTaEi	AARAHdVAAI	ALGRksAcLN	FADs
ap2{bncbf2}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WpGTfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf3}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WpGTfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf1}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf6}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf8}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bocbf3}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{brcbf2}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf4}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{bncbf5}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{bocbf4}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf1}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf3}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf4}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf5}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf6}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{brcbf7}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{rscbf1}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{bocbf1}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEi	AARAHdVAAI	ALGRgAcLN	FADs
ap2{bocbf5}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WlGTFfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{bncbf9}	HpiYrGVRlR .ksgkwwcEv	REpNkK.sRI	WpGTfLTaEm	AARAHdVAAI	ALGRgAcLN	YADs
ap2{zmcbf1}	HPvYrGVRrR gpagrWcEv	REpNkK.sRI	WlGTFaTpea	AARAHdVAAI	ALGRaAcLN	FADs
ap2{gmcbf1}	HpiYsGVR.R rntdkWvEv	REpNkK.tRI	WlGTFpTpEm	AARAHdVAAm	ALGRyAcLN	FADs
Consensus	HP-Y-GVR-R	-----WV-E-RE-NK-----RI	W-GTF-T-E-	AARAHdVAA-	ALRG--A-LN	-ADs
ap2{tiny}	HPvYrGVRKR .nwgkwwaEi	REpRkK.sRI	WlGTFpSpEm	AARAHdVAAI	sikGaaAiLN	fpDI
Consensus	-----	-----N-----	-----T-----	-----	ALR-----	-A-S
ap2{tiny}	-----	-----	-----	-----	-----	-----

FIGURE 19E

1	50	64
ap2{atcbf2}	HPiYrGVRqR .nsgkwwcEl REPnKk.tRI WLGTfqtEm AARAHdVAaI ALRGsACLN fADS	
ap2{atcbf3}	HPiYrGVRrR .nsgkwwcEv REPnKk.tRI WLGTfqtEm AARAHdVAaI ALRGsACLN fADS	
ap2{atcbf1}	HPiYrGVRqR .nsgkwwsEv REPnKk.tRI WLGTfqtEm AARAHdVAaI ALRGsACLN fADS	
ap2{bjcbf4}	HPiYrGVRqR .nsgkwwcEv REPnKk.sRI WLGTfPtEm AARAHdVAaI ALRGsACLN fADS	
ap2{bocbf2}	HPvYrGVRlR .nsgkwwcEv REPnKk.sRI WLGTfLtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{rscbf2}	HPiYrGVRlR .nsgkwwcEv REPnKk.sRI WLGTfLtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bjcbf2}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfLtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bncbf7}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfLtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bjcbf3}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WpGTfLtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bncbf2}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WpGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf3}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WpGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf1}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf6}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf8}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bocbf3}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{brcbf2}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf4}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bncbf5}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bocbf4}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf1}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf3}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf4}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf5}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf6}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{brcbf7}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{rscbf1}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bocbf1}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtAEI AARAHdVAaI ALRGsACLN fADS	
ap2{bocbf5}	HPiYrGVRlR .ksgkwwcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{bncbf9}	HPvYrGVRlR .ksgkwwcEv REPnKk.sRI WpGTfKtEm AARAHdVAaI ALRGsACLN YADS	
ap2{zncbf1}	HPvYrGVRrR gpagrWVcEv REPnKk.sRI WLGTfKtEm AARAHdVAaI ALRGsACLN fADS	
ap2{gmcbf1}	HPiYsGVR.R rntdkWVsEv REPnKk.tRI WLGTfPtEm AARAHdVAaI ALRGsACLN fADS	
ap2{dreb2a}	rcsfrGVRqR i.wgkWWaEi REPnrg.sRI WLGTfPtEm AAsAyDeaAk AmyGslarLN fprS	
ap2{dreb2b}	hcsfrGVRqR i.wgkWWaEi REPkig.tRI WLGTfPtEm AAsAyDeaAk AmyGslarLN fpqS	
Consensus	-----GVR-R -----WV-E- RE-----RI W-GTF-T--- AA-A-D-AA- A--G--A-LN ---S	
ap2{tiny}	hpvyYrGVRkr .nwgkWWaEi REprk.sRI WLGTfPpEm AArAhDVaAl sikGasAiLN fpdl	
Consensus	-----T-----A-----S	
- ap2{tiny}		

## FIGURE 20

	1				49
n{brCBF3N}	MnsvstfsEl	lcSenespv	te.GgdY..i	LAaSCP	GRKKFqETR
n{brCBF6N}	MnsvstfsEl	lcSenkspv	te.GgdY..i	LAaSCP	GRKKFqETR
n{bnCBF5N}	MnsvstfsEl	lrSenespv	te.GgdY..i	LAaSCP	GRKKFqETR
n{atCBF2N}	MnsfsafsEm	fgSdyespv	s..GgdYspk	LAaSCP	GRKKFqETR
n{atCBF3N}	MnsfsafsEm	fgSdyessvs	s..GgdYipt	LAaSCP	GRKKFqETR
n{atCBF1N}	MnsfsafsEm	fgSdye...p	q..GgdYcpt	LAaSCP	GRKKFqETR
n{bnCBF2N}	MntfpastEm	vgSenespvt	tvvGgdYypm	LAaSCP	GRKKFqETR
n{bnCBF6N}	MntfpastEm	vgSenespvt	tvvGgdYypm	LAaSCP	GRKKFqETR
n{boCBF3N}	MntfpastEm	vgSenespvt	tvvGgdYypm	LAaSCP	GRKKFqETR
n{bnCBF3N}	MntfpastEm	vgSenespvt	tvaGgdYypm	LAaSCP	GRKKFqETR
n{bnCBF8N}	MntfpastEm	vgSenespvt	tvaGgdYypm	LAaSCP	GRKKFqETR
n{bnCBF9N}	MntfpastEm	vgSenespvt	tvaGgdYypm	LAaSCP	GRKKFqETR
n{brCBF2N}	MntfpastEm	vgSenespvt	tvaGgdYypm	LAaSCP	GRKKFqETR
n{boCBF5N}	MntfpastEm	vsSenespvt	tvvGgdYypm	LAaSCP	GRKKFqETR
n{boCBF2N}	MtsfstfsEl	lgSehespvt	..lGeeYcpk	LAaSCP	GRKKFqETR
n{rsCBF2N}	MttfstfsEm	lgSeyespvt	..lGeeYcpk	LAaSCP	GRKKFqETR
n{bnCBF4N}	MnsvstfsEl	lgSenesp..	..vGgdYcpm	LAaSCP	GRKKFqETR
n{bnCBF7N}	MnsvstfsEl	lgSenesp..	..vGgdYcpm	LAaSCP	GRKKFqETR
n{boCBF4N}	MnsvstfsEl	lgSenesp..	..vGgdYcpm	LAaSCP	GRKKFqETR
n{brCBF7N}	MnsvstfsEl	lgSqnesp..	..vGgdYcpm	LAaSCP	GRKKFqETR
n{brCBF4N}	MdsistfpEl	lgSenespvt	tvvGgdYcpr	LAaSCP	GRKKFqETR
n{brCBF5N}	MdsistfpEl	lgSenespvt	tvvGgdYcpr	LAaSCP	GRKKFqETR
n{rsCBF1N}	MdsistfsEl	lgSenespvt	tvvGgdYfpr	LAaSCP	GRKKFqETR
Consensus	M-----E-	--S-----	---G--Y---	LA-SCP	GRKKF-ETR



FIGURE 21A

1					50
c{bnCBF3C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvtmqngln	meettavasq
c{bnCBF9C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvtmqngln	meettavasq
c{brCBF2C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvtmqngln	meemtavasq
c{bnCBF1C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvtmqngqn	meettavasq
c{bnCBF8C}	AsRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvt.....	meetmavasq
c{bnCBF6C}	AsRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvt.....	meetmavasq
c{boCBF3C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvt.....	meetmavasq
c{bnCBF2C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdvtmqngln	meettavasq
c{boCBF5C}	AwRLRIpEtT	ChKdIQKAAA	EAAlaFeaek	sdatmqngln	meettaaasq
c{brCBF5C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	s.....	.....
c{rsCBF1C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	s.....	.....
c{bnCBF4C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	sd.t.....	.....
c{boCBF4C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	sd.t.....	.....
c{bnCBF5C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	sd.t.....	.....
c{brCBF7C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	sdtt.....	.....
c{brCBF6C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFeaek	sd.t.....	.....
c{boCBF1C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaFgaek	sd.t.....	.....
c{bjCBF2C}	AwRLRIpEtT	CpKeIQKAAA	EAavaF....	.....	.....
c{bjCBF3C}	AwRLRIpEtT	CpKeIQKAAA	EAavaF....	.....	.....
c{bnCBF7C}	AwRLRIpEtT	CpKeIQKAAA	EAavaF....	.....	.....
c{rsCBF2C}	AwRLRIpEtT	CpKdIQKAAA	EAavaF....	.....	.....
c{atCBF1C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaF....	.....	.....
c{atCBF2C}	AwRLRIpEtT	CaKeIQKAAA	EAAlnF....	.....	.....
c{atCBF3C}	AwRLRIpEtT	CaKdIQKAAA	EAAlaF....	.....	.....
Consensus	A-RLRI-E-T	C-K-IQKAAA	EAA--F----	-----	-----
51					100
c{bnCBF3C}	aevndttteh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvftgE
c{bnCBF9C}	aevndttteh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvftgE
c{brCBF2C}	aevndttteh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{bnCBF1C}	aevndttteh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvftgE
c{bnCBF8C}	aevndtttdh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvftgE
c{bnCBF6C}	aevndtttdh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{boCBF3C}	aevndtttdh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{bnCBF2C}	aevndttteh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{boCBF5C}	tevsdtttdh	gmnmeeatav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{brCBF5C}	.....dh	gmnmknttav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{rsCBF1C}	.....dh	gmnmknttav	asqaEvndtt	td.HgvDmEE	TmVEAvfteE
c{bnCBF4C}	.....ttndh	gmnm.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{boCBF4C}	.....ttndh	gmnm.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{bnCBF5C}	.....ttndh	gmnm.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{brCBF7C}	.....ttndh	gmnm.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{brCBF6C}	.....ttndr	gmnmeeatav	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{boCBF1C}	.....ttndq	gmnmeeatav	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{bjCBF2C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{bjCBF3C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{bnCBF7C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{rsCBF2C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{atCBF1C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{atCBF2C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
c{atCBF3C}	.....	.....	asqaEvndtt	td.HgldmEE	TmVEAvfteE
Consensus	-----	-----	-----	-----	-----

## FIGURE 21A-continued

	101				150
c{bnCBF3C}	qsegfnmake	stveaavvte	epskgsYMD.	eEwmleMptl	ladmAeGMLl
c{bnCBF9C}	qsegfnmake	stveaavvte	epskgsYMD.	eEwmleMptl	ladmAeGMLl
c{brCBF2C}	qsegfnmake	stveaavvte	epskgsYMD.	eEwmleMptl	ladmAeGMLl
c{bnCBF1C}	qsegfnmake	stveaavvte	epskgsYMD.	eEwmleMptl	ladmAeGMLl
c{bnCBF8C}	qsegfnmake	stveaavvte	epskgsYMD.	eEwmleMptl	ladmAeGMLl
c{bnCBF6C}	qsegfnmaee	stveaavvtd	elskgyfYMD.	eEwtyeMptl	ladmAaGMLl
c{boCBF3C}	qsegfnmaee	stveaavvtd	elskgyfYMD.	eEwtyeMptl	ladmAaGMLl
c{bnCBF2C}	qsegfnmaee	stveaavvtd	elskgyfYMD.	eEwtyeMptl	ladmAaGMLl
c{boCBF5C}	qsegfnmake	stveaavvte	elskgyfYMD.	eEwtyeMptl	ladmAaGMLl
c{brCBF5C}	qregfymtee	trvegvttee	q.nnwfYMD.	eEwmfgMptl	lvdmAeGMLl
c{rsCBF1C}	qregfymtee	trvegvttee	q.nnwfYMD.	eEwmfgMptl	lvdmAeGMLl
c{bnCBF4C}	qrdgfymaee	ttvegvtpee	qmskgyfYMD.	eEwmfgMptl	ladmAaGMLl
c{boCBF4C}	qrdgfymaee	ttvegvtpee	qmskgyfYMD.	eEwmfgMptl	ladmAaGMLl
c{bnCBF5C}	qregfymaee	ttvvgvtpee	qmskgyfYMD.	eEwmfgMptl	ladmAaGMLl
c{brCBF7C}	qregfymaee	ttvegvtpee	qmskgyfYMD.	eEwmfgMptl	ladmAaGMLl
c{brCBF6C}	qregfymaee	ttvegvtpee	qmskgyfYMD.	eEwtfeMprl	ladmAeGMLl
c{boCBF1C}	qregfylaee	ttvegvttee	q.skgyfYMD.	eEwtfgMqsf	ladmAeGMLf
c{bjCBF2C}	esse.....	.....	....gfYMD.	eEfmfgMptl	wasmAeGMLl
c{bjCBF3C}	esse.....	.....	....gfYMa.	eEfmfgMptl	wasvAeGMLl
c{bnCBF7C}	ennd.....	.....	....gfYMD.	eEsmfgMpal	lasmAeGMLl
c{rsCBF2C}	vnnd.....	.....	....efYMD.	eEsmfgMpsl	lasmAeGMLl
c{atCBF1C}	qseg.....	.....	....afYMD.	eEtmfgMptl	ldnmAeGMLl
c{atCBF2C}	qsqd.....	.....	....afYMD.	eEamlgMssl	ldnmAeGMLl
c{atCBF3C}	qsen.....	.....	....afYMH.	dEamfeMpsl	lanmAeGMLl
Consensus	-----	-----	-----YM--	-E-----M--	-----A-GML-

## FIGURE 21B

	1					50
C{CbnCBF3}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvTmqngln	MEEttAvASQ	
C{CbnCBF9}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvTmqngln	MEEttAvASQ	
C{CbrCBF2}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvTmqngln	MEEttAvASQ	
C{CbnCBF1}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvTmqngqn	MEEttAvASQ	
C{CbnCBF8}	AsRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvT.....	MEEtmAvASQ	
C{CbnCBF6}	AsRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvT.....	MEEtmAvASQ	
C{CboCBF3}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvT.....	MEEtmAvASQ	
C{CbnCBF2}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDvTmqngln	MEEttAvASQ	
C{CboCBF5}	AwRLRIPETT	CHKDIQKAAA	EAALAFEAEK	SDaTmqngln	MEEttAaASQ	
Consensus	A-RLRIPETT	CHKDIQKAAA	EAALAFEAEK	SD-T-----	MEE--A-ASQ	

	51					100
C{CbnCBF3}	aEVnDTTTeH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbnCBF9}	aEVnDTTTeH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbrCBF2}	aEVnDTTTeH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbnCBF1}	aEVnDTTTeH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbnCBF8}	aEVnDTTtdH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbnCBF6}	aEVnDTTtdH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CboCBF3}	aEVnDTTtdH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CbnCBF2}	aEVnDTTTeH	GMNMEEaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
C{CboCBF5}	tEVsDTTtdH	GMNMEETaTAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFTgEQ	
Consensus	-EV-DTTT-H	GMNMEET-TAV	ASQAEVNDTT	TDHGVDMEET	MVEAVFT-RQ	

	101					150
C{CbnCBF3}	SEGFNMAKES	TvEAAVVTeE	PSKGsYMDEE	Wm1EMPTLLA	DMAeGMLLpp	
C{CbnCBF9}	SEGFNMAKES	TvEAAVVTeE	PSKGsYMDEE	Wm1EMPTLLA	DMAeGMLLpp	
C{CbrCBF2}	SEGFNMAKES	TvEAAVVTeE	PSKGsYMDEE	Wm1EMPTLLA	DMAeGMLLpp	
C{CbnCBF1}	SEGFNMAKES	TvEAAVVTeE	PSKGsYMDEE	Wm1EMPTLLA	DMAeGMLLpp	
C{CbnCBF8}	SEGFNMAKES	TvEAAVVTeE	PSKGsYMDEE	Wm1EMPTLLA	DMAeGMLL--	
C{CbnCBF6}	SEGFNMAeES	TvEAAVVTeE	1SKGfYMDEE	WtyEMPTLLA	DMAaGMLLpp	
C{CboCBF3}	SEGFNMAeES	TvEAAVVTeE	1SKGfYMDEE	WtyEMPTLLA	DMAaGMLLpp	
C{CbnCBF2}	SEGFNMAeES	TvEAAVVTeE	1SKGfYMDEE	WtyEMPTLLA	DMAaGMLLpp	
C{CboCBF5}	SEGFNMAKES	TaEAAVVTeE	1SKGvYMDEE	WtyEMPTLLA	DMAaGMLLpp	
Consensus	SEGFNMA-ES	T-EAAVVTeE	-SKG-YMDEE	W--EMPTLLA	DMA-GMLL--	

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Michael Thomashow, Eric Stockinger, Kirsten Jaglo-Ottosen, Sarah Gilmour, Daniel Zarka and Cai-Zhong Jiang

(ii) TITLE OF INVENTION: PLANT HAVING ALTERED ENVIRONMENTAL STRESS TOLERANCE

(iii) NUMBER OF SEQUENCES: 100

(A) ADDRESSEE: David J. Weitz,  
Wilson Sonsini Goodrich & Rosati  
(B) STREET: 650 Page Mill Road  
(C) CITY: Palo Alto  
(D) STATE: California  
(E) COUNTRY: USA  
(F) ZIP: 94304-1050

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch diskette  
(B) COMPUTER: IBM compatible  
(C) OPERATING SYSTEM: Microsoft Windows 95  
(D) SOFTWARE: MS WORD 97,  
ASCII (DOS) TEXT format

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/018,233  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/017,816  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/018,235  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/017,575  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/018,227  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/018,234  
(B) FILING DATE: February 3, 1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/ 198,119  
(B) FILING DATE: November 23, 1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: David J. Weitz  
(B) REGISTRATION NUMBER: 38,362  
(C) REFERENCE/DOCKET NUMBER: 19117.714

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (650) 493-9300  
(B) TELEFAX: (650) 493-6811  
(C) TELEX: None

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 905  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: CBF1 gene  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AAAAAGAATC TACCTGAAAA GAAAAAAAAG AGAGAGAGAT ATAAATAGCT   50
TACCAAGACA GATATACTAT CTTTTATTAA TCCAAAAAGA CTGAGAACTC  100
TAGTAACTAC GTACTACTTA AACCTTATCC AGTTTCTTGA AACAGAGTAC  150
TCTGATCAAT GAACTCATTT TCAGCTTTTT CTGAAATGTT TGGCTCCGAT  200
TACGAGCCTC AAGGCGGAGA TTATTGTCCG ACGTTGGCCA CGAGTTGTCC  250
GAAGAAACCG GCGGGCCGTA AGAAGTTTCG TGAGACTCGT CACCCAATTT  300
```

```

ACAGAGGAGT TCGTCAAAGA AACTCCGGTA AGTGGGTTTC TGAAGTGAGA 350
GAGCCAAACA AGAAAACCAG GATTTGGCTC GGGACTTTCC AAACCGCTGA 400
GATGGCAGCT CGTGCTCACG ACGTCGCTGC ATTAGCCCTC CGTGGCCGAT 450
CAGCATGTCT CAACTTCGCT GACTCGGCTT GGCGGCTACG AATCCCGGAG 500
TCAACATGCG CCAAGGATAT CCAAAAAGCG GCTGCTGAAG CGGCGTTGGC 550
TTTTCAAGAT GAGACGTGTG ATACGACGAC CACGGATCAT GGCCTGGACA 600
TGGAGGAGAC GATGGTGGAA GCTATTTATA CACCGGAACA GAGCGAAGGT 650
GCGTTTTATA TGGATGAGGA GACAATGTTT GGGATGCCGA CTTTGTTGGA 700
TAATATGGCT GAAGGCATGC TTTTACCGCC GCCGTCTGTT CAATGGAATC 750
ATAATTATGA CGGCGAAGGA GATGGTGACG TGTCGCTTTG GAGTTACTAA 800
TATTGCATAG TCGTTTCCAT TTTTGTACTA TAGTTTGAAA ATATTCTAGT 850
TCCTTTTTTA GAATGGTTCC TTCATTTTAT TTTATTTTAT TGTGTAGAA 900
ACGAG 905

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: CBF1 protein
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asn	Ser	Phe	Ser	Ala	Phe	Ser	Glu	Met	Phe	Gly	Ser	Asp	Tyr	5	10	15
Glu	Pro	Gln	Gly	Gly	Asp	Tyr	Cys	Pro	Thr	Leu	Ala	Thr	Ser	Cys	20	25	30
Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	Thr	Arg	His	35	40	45
Pro	Ile	Tyr	Arg	Gly	Val	Arg	Gln	Arg	Asn	Ser	Gly	Lys	Trp	Val	50	55	60
Ser	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Thr	Arg	Ile	Trp	Leu	Gly	65	70	75
Thr	Phe	Gln	Thr	Ala	Glu	Met	Ala	Ala	Arg	Ala	His	Asp	Val	Ala	80	85	90
Ala	Leu	Ala	Leu	Arg	Gly	Arg	Ser	Ala	Cys	Leu	Asn	Phe	Ala	Asp	95	100	105
Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Ser	Thr	Cys	Ala	Lys	Asp	110	115	120
Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Gln	Asp	Glu	125	130	135
Thr	Cys	Asp	Thr	Thr	Thr	Thr	Asp	His	Gly	Leu	Asp	Met	Glu	Glu	140	145	150
Thr	Met	Val	Glu	Ala	Ile	Tyr	Thr	Pro	Glu	Gln	Ser	Glu	Gly	Ala	155	160	165
Phe	Tyr	Met	Asp	Glu	Glu	Thr	Met	Phe	Gly	Met	Pro	Thr	Leu	Leu	170	175	180
Asp	Asn	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Pro	Pro	Ser	Val	Gln	185	190	195
Trp	Asn	His	Asn	Tyr	Asp	Gly	Glu	Gly	Asp	Gly	Asp	Val	Ser	Leu	200	205	210
Trp	Ser	Tyr															

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE: .

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: N/A - Synthetic
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A  
(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GATCATTTC A TGGCCGACCT GCTTTTT

27

(3) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: N/A - Synthetic  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CACAATTTC A AGAATTCAC T GCTTTTT

28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27



(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: N/A - Synthetic  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:  
GATCATTTCA TGGTATGTCT GCTTTTT

27

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: N/A - Synthetic  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A

(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GATCATTTC A TGG AATCACT GCTTTTT

27

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27  
(B) TYPE: Nucleic Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: N/A - Synthetic  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATCACTTGA TGGCCGACCT CTTTTTT

27

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27
  - (B) TYPE: Nucleic Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: N/A - Synthetic
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATCAATATA CTACCGACAT GAGTTCT

27

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25
  - (B) TYPE: Nucleic Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: N/A - Synthetic
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A

(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Table 1

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACTACCGACA TGAGTTCCAA AAAGC

25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60  
(B) TYPE: Amino Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Amino Acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Arabidopsis thaliana  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Figure 2D

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ile	Tyr	Arg	Gly	Val	Arg	Gln	Arg	Asn	Ser	Gly	Lys	Trp	Val	Ser
				5					10					15
Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Thr	Arg	Ile	Trp	Leu	Gly	Thr
				20					25					30
Phe	Gln	Thr	Ala	Glu	Met	Ala	Ala	Arg	Ala	His	Asp	Val	Ala	Ala
				35					40					45
Leu	Ala	Leu	Arg	Gly	Arg	Ser	Ala	Cys	Leu	Asn	Phe	Ala	Asp	Ser
				50					55					60

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Polypeptide

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Tobacco
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION: Figure 2D

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

His	Tyr	Arg	Gly	Val	Arg	Gln	Arg	Pro	Trp	Gly	Lys	Phe	Ala	Ala
				5					10					15
Glu	Ile	Arg	Asp	Pro	Ala	Lys	Asn	Gly	Ala	Arg	Val	Trp	Leu	Gly
				20					25					30
Thr	Tyr	Glu	Thr	Ala	Glu	Glu	Ala	Ala	Leu	Ala	Tyr	Asp	Lys	Ala
				35					40					45
Ala	Tyr	Arg	Met	Arg	Gly	Ser	Lys	Ala	Leu	Leu	Asn	Phe	Pro	His
				50					55					60
Arg														

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 651
  - (B) TYPE: Nucleic Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Arabidopsis thaliana*
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY: CBF2
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
- |   |     |
|---|-----|
| ATGAACTCAT TTTCTGCCTT TTCTGAAATG TTTGGCTCCG ATTACGAGTC  | 50  |
| TCCGGTTTCC TCAGGCGGTG ATTACAGTCC GAAGCTTGCC ACGAGCTGCC  | 100 |
| CCAAGAAACC AGCGGGAAGG AAGAAGTTTC GTGAGACTCG TCACCCAATT  | 150 |
| TACAGAGGAG TTCGTCAAAG AAACCTCCGGT AAGTGGGTGT GTGAGTTGAG | 200 |
| AGAGCCAAAC AAGAAAACGA GGATTTGGCT CGGGACTTTC CAAACCGCTG  | 250 |
| AGATGGCAGC TCGTGCTCAC GACGTCGCCG CCATAGCTCT CCGTGGCAGA  | 300 |
| TCTGCCTGTC TCAATTTTCGC TGA CTGGCT TGGCGGCTAC GAATCCCGGA | 350 |
| ATCAACCTGT GCCAAGGAAA TCCAAAAGGC GGCGGCTGAA GCCGCGTTGA  | 400 |
| ATTTTCAAGA TGAGATGTGT CATATGACGA CGGATGCTCA TGGTCTTGAC  | 450 |
| ATGGAGGAGA CCTTGGTGGA GGCTATTTAT ACGCCGGAAC AGAGCCAAGA  | 500 |
| TGCGTTTTTAT ATGGATGAAG AGGCGATGTT GGGGATGTCT AGTTTGTTGG | 550 |
| ATAACATGGC CGAAGGGATG CTTTTACCGT CGCCGTCGGT TCAATGGAAC  | 600 |
| TATAATTTTG ATGTCGAGGG AGATGATGAC GTGTCCTTAT GGAGCTATTA  | 650 |

A

651

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 216
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Arabidopsis thaliana*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: CBF2
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Asn	Ser	Phe	Ser	Ala	Phe	Ser	Glu	Met	Phe	Gly	Ser	Asp	Tyr	5	10	15
Glu	Ser	Pro	Val	Ser	Ser	Gly	Gly	Asp	Pyr	Ser	Pro	Lys	Leu	Als	20	25	30
Thr	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	35	40	45
Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Gln	Arg	Asn	Ser	Gly	50	55	60
Lys	Trp	Val	Cys	Glu	Leu	Arg	Glu	Pro	Asn	Lys	Lys	Thr	Arg	Ile	65	70	75
Trp	Leu	Gly	Thr	Phe	Gln	Thr	Ala	Glu	Met	Ala	Ala	Arg	Ala	His	80	85	90
Asp	Val	Ala	Ala	Ile	Ala	Leu	Arg	Gly	Arg	Ser	Ala	Cys	Leu	Asn	95	100	105
Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Ser	Thr	Cys	110	115	120
Ala	Lys	Glu	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Asn	Phe	125	130	135
Gln	Asp	Glu	Met	Cys	His	Met	Thr	Thr	Asp	Ala	His	Gly	Leu	Asp	140	145	150

12/101

Met	Glu	Glu	Thr	Leu	Val	Glu	Ala	Ile	Tyr	Thr	Pro	Glu	Gln	Ser
				155					160					165
Gln	Asp	Ala	Phe	Tyr	Met	Asp	Glu	Glu	Ala	Met	Leu	Gly	met	Ser
				170					175					180
Ser	Leu	Leu	Asp	Asn	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Ser	Pro
				185					190					195
Ser	Val	Gln	Trp	Asn	Tyr	Asn	Phe	Asp	Val	Glu	Gly	Asp	Asp	Asp
				200					205					210
Val	Ser	Leu	Trp	Ser	Tyr									
				215										

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 651
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Arabidopsis thaliana
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: CBF3
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:



ATGAACTCAT TTTCTGCTTT TTCTGAAATG TTTGGCTCCG ATTACGAGTC 50  
TTCGGTTTTCC TCAGGCGGTG ATTATATTCC GACGCTTGCG AGCAGCTGCC 100  
CCAAGAAACC GGC GGCTCGT AAGAAGTTTC GTGAGACTCG TCACCCAATA 150  
TACAGAGGAG TTCGTCGGAG AAACCTCCGGT AAGTGGGTTT GTGAGGTTAG 200  
AGAACCAAAC AAGAAAACAA GGATTTGGCT CGGAACATTT CAAACCGCTG 250  
AGATGGCAGC TCGAGCTCAC GACGTTGCCG CTTTAGCCCT TCGTGGCCGA 300  
TCAGCCTGTC TCAATTTTCG TGAATCGGCT TGGAGACTCC GAATCCCGGA 350  
ATCAACTTGC GCTAAGGACA TCCAAAAGGC GGC GGCTGAA GCTGCGTTGG 400  
CGTTTCAGGA TGAGATGTGT GATGCGACGA CGGATCATGG CTTGACATG 450  
GAGGAGACGT TGGTGGAGGC TATTTACACG GCGGAACAGA GCGAAAATGC 500  
GTTTTATATG CACGATGAGG CGATGTTTGA GATGCCGAGT TTGTTGGCTA 550  
ATATGGCAGA AGGGATGCTT TTGCCGCTTC CGTCCGTACA GTGGAATCAT 600  
AATCATGAAG TCGACGGCGA TGATGACGAC GTATCGTTAT GGAGTTATTA 650  
A 651

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 216
  - (B) TYPE: Amino Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Polypeptide
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Arabidopsis thaliana*
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY: CBF3
  - (B) LOCATION:

(C) IDENTIFICATION METHOD:  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met	Asn	Ser	Phe	Ser	Ala	Phe	Ser	Glu	Met	Phe	Gly	Ser	Asp	Tyr	5	10	15
Glu	Ser	Ser	Val	Ser	Ser	Gly	Gly	Asp	Tyr	Ile	Pro	Thr	Leu	Ala	20	25	30
Ser	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	35	40	45
Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Arg	Arg	Asn	Ser	Gly	50	55	60
Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Thr	Arg	Ile	65	70	75
Trp	Leu	Gly	Thr	Phe	Gln	Thr	Ala	Glu	Met	Ala	Ala	Arg	Ala	His	80	85	90
Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Ser	Ala	Cys	Leu	Asn	95	100	105
Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Ser	Thr	Cys	110	115	120
Ala	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	125	130	135
Gln	Asp	Glu	Met	Cys	Asp	Ala	Thr	Thr	Asp	His	Gly	Phe	Asp	Met	140	145	150
Glu	Glu	Thr	Leu	Val	Glu	Ala	Ile	Tyr	Thr	Ala	Glu	Gln	Ser	Glu	155	160	165
Asn	Ala	Phe	Tyr	Met	His	Asp	Glu	Ala	Met	Phe	Glu	Met	Pro	Ser	170	175	180
Leu	Leu	Ala	Asn	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Leu	Pro	Ser	185	190	195
Val	Gln	Trp	Asn	His	Asn	His	Glu	Val	Asp	Gly	Asp	Asp	Asp	Asp	200	205	210
Val	Ser	Leu	Trp	Ser	Tyr										215		

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Tobacco

(B) STRAIN:

(C) INDIVIDUAL ISOLATE: N/A

(D) DEVELOPMENTAL STAGE: N/A

(E) HAPLOTYPE: N/A

(F) TISSUE TYPE: N/A

(G) CELL TYPE: N/A

(H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION: Figure 2D

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTGGCGGCTA CGAATCCC

18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 210  
(B) TYPE: Amino Acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Polypeptide

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No  
(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Brassica Napus  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	Trp
					5					10				15
Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	Leu
					20					25				30
Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	Ala	His	Asp	Val

16/101

	35		40		45
Ala	Ala	Leu	Ala	Leu	Arg
Ala	Leu	Arg	Gly	Arg	Gly
Ala	Cys	Leu	Asn	Tyr	Ala
50	50	60			
Asp	Ser	Ala	Trp	Arg	Ile
Asp	Ile	Gln	Lys	Ala	Ala
Ala	Ala	Glu	Ala	Ala	Leu
Ala	Leu	Ala	Phe	Glu	Ala
80	85	90			
Glu	Lys	Ser	Asp	Val	Thr
Glu	Thr	Met	Gln	Asn	Gly
Gln	Asn	Met	Glu	Glu	
95	100	115			
Thr	Thr	Ala	Val	Ala	Glu
Val	Ser	Gln	Ala	Glu	Val
Val	Asn	Asp	Thr	Thr	Thr
110	115	120			
Glu	His	Gly	Met	Asn	Met
Glu	Glu	Ala	Thr	Ala	Val
Ala	Val	Ala	Ser	Gln	
125	130	135			
Ala	Glu	Val	Asn	Asp	Thr
Thr	Thr	Thr	Asp	His	Gly
Val	Asp	Met	Glu		
140	145	150			
Glu	Thr	Met	Val	Glu	Ala
Val	Phe	Thr	Gly	Glu	Gln
Ser	Glu	Gly			
155	160	165			
Phe	Asn	Met	Ala	Lys	Glu
Ser	Thr	Val	Glu	Ala	Ala
Val	Val	Thr			
170	175	180			
Glu	Glu	Pro	Ser	Lys	Gly
Ser	Tyr	Met	Asp	Glu	Glu
Trp	Met	Leu			
185	190	195			
Glu	Met	Pro	Thr	Leu	Leu
Ala	Ala	Asp	Met	Ala	Glu
Gly	Met	Leu	Leu		
200	205	210			

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 632
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Canola
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CACCCGATAT ACCGGGGAGT TCGTCTGAGA AAGTCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAACCAAACA AGAAATCTAG AATTTGGCTT GGAACTTTCA	100
AAACAGCTGA GATGGCAGCT CGTGCTCACG ACGTCGCTGC CCTAGCCCTC	150
CGTGGAAGAG GCGCCTGCCT CAATTATGCG GACTCGGCTT GGCGGCTCCG	200
CATCCCGGAG ACAACCTGCC ACAAGGATAT CCAGAAGGCT GCTGCTGAAG	250
CCGCATTGGC TTTTGAGGCT GAGAAAAGTG ATGTGACGAT GCAAAATGGC	300
CAGAACATGG AGGAGACGAC GGCGGTGGCT TCTCAGGCTG AAGTGAATGA	350
CACGACGACA GAACATGGCA TGAACATGGA GGAGGCAACG GCAGTGGCTT	400
CTCAGGCTGA GGTGAATGAC ACGACGACGG ATCATGGCGT AGACATGGAG	450
GAGACAATGG TGGAGGCTGT TTTTACTGGG GAACAAAGTG AAGGGTTTAA	500
CATGGCGAAG GAGTCGACGG TGGAGGCTGC TGTTGTTACG GAGGAACCGA	550
GCAAAGGATC TTACATGGAC GAGGAGTGGA TGCTCGAGAT GCCGACCTTG	600
TTGGCTGATA TGGCAGAAGG GATGCTCCTG CC	632

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:

## (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCCCAAGCTT CAAGTTTAGT GAGCACTATG TGCTCG

36

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGAAGATCTC CTTCCCAGAA ACAACACAAT CTAC

34

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

19/101

- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A

- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A

- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:

- (x) PUBLICATION INFORMATION:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCCCAAGCTT GTTTCATTTT CTCCATGAAG GAGAT

35

- (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: No

- (iv) ANTI-SENSE: No

- (v) FRAGMENT TYPE:

- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A

- (vii) IMMEDIATE SOURCE: N/A

20/101

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGAAGATCTT ATCGTCGTCG TCGTCTACCA AAACCACAC

39

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCTCTAAGCT TCACAAGGGG TTCGTTTGGT GC

32

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40
- (B) TYPE: Nucleic Acid Sequence



(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: No  
(iv) ANTI-SENSE: No  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:  
    (A) ORGANISM:  
    (B) STRAIN:  
    (C) INDIVIDUAL ISOLATE: N/A  
    (D) DEVELOPMENTAL STAGE: N/A  
    (E) HAPLOTYPE: N/A  
    (F) TISSUE TYPE: N/A  
    (G) CELL TYPE: N/A  
    (H) CELL LINE: N/A  
    (I) ORGANELLE: N/A  
(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A  
(ix) FEATURE:  
    (A) NAME/KEY:  
    (B) LOCATION:  
    (C) IDENTIFICATION METHOD:  
    (D) OTHER INFORMATION:  
(x) PUBLICATION INFORMATION:  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:  
GGGGTACCTT TTGGGAGTTG GAATAGAAAT GGGTTTGATG

40

(2) INFORMATION FOR SEQ ID NO:25:  
(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 36  
    (B) TYPE: Nucleic Acid Sequence  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: No  
(iv) ANTI-SENSE: No  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:  
    (A) ORGANISM:  
    (B) STRAIN:  
    (C) INDIVIDUAL ISOLATE: N/A  
    (D) DEVELOPMENTAL STAGE: N/A  
    (E) HAPLOTYPE: N/A  
    (F) TISSUE TYPE: N/A  
    (G) CELL TYPE: N/A

22/101

(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCCCAAGCTT AATTTTACTC AAAATGTTTT GGTGTC

36

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 44  
(B) TYPE: Nucleic Acid Sequence  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CCGGTACCTT TCCAAAGATT TTTTCTTTC CAATAGAAGT AATC

44

(2) INFORMATION FOR SEQ ID NO:27:

23/101

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - ORGANISM:
  - STRAIN:
    - (C) INDIVIDUAL ISOLATE: N/A
    - (D) DEVELOPMENTAL STAGE: N/A
    - (E) HAPLOTYPE: N/A
    - (F) TISSUE TYPE: N/A
    - (G) CELL TYPE: N/A
    - (H) CELL LINE: N/A
    - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCGGAAGCTT CATTTTCTGC TACAGAAGTG

30

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A

(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CCGGTACCTT TCCAAAGCTG TGTTTTCTCT TTTTCAAGTG

40

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42  
(B) TYPE: Nucleic Acid Sequence  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GCCCAAGCTT CAAATTCTGA ATATTCACAT ATCAAAAAAG TG

42

## (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 40
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGAAGATCTG TTCTTCTTGT CTTAAGCAAA CACTTTGAGC

40

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 41
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:

26/101

(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCCCAAGCTT TCGTCTGTGA TCATACAAGG CACAAAACGA C

41

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42  
(B) TYPE: Nucleic Acid Sequence  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

27/101

GGAAGATCTA GTTAATCTTG ATTTGATTAA AAGTTTATAT AG

42

## (2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A

- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A

- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:

- (x) PUBLICATION INFORMATION:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CAAACCTCAGT AGGATTCTGG TGTGT

25

## (2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: No

- (iv) ANTI-SENSE: No

- (v) FRAGMENT TYPE:

- (vi) ORIGINAL SOURCE:

28/101

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

- (ix) FEATURE:
- (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGAAGATCTT GAAACAGAGT ACTCTGATCA ATGAACTC

38

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

- (ix) FEATURE:
- (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

29/101



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CGCGGATCCC TCGTTTCTAC AACATAAAA TAAAATAAAA TG

42

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37

(B) TYPE: Nucleic Acid Sequence

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM:

(B) STRAIN:

(C) INDIVIDUAL ISOLATE: N/A

(D) DEVELOPMENTAL STAGE: N/A

(E) HAPLOTYPE: N/A

(F) TISSUE TYPE: N/A

(G) CELL TYPE: N/A

(H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGGGTACCTG AAACAGAGTA CTCTGATCAA TGAAGTC

37

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41

(B) TYPE: Nucleic Acid Sequence

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

30/101

- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
- GCTCTAGACT CGTTTCTACA ACAATAAAAT AAAATAAAAT G 41
- (2) INFORMATION FOR SEQ ID NO:38:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 577
  - (B) TYPE: Nucleic Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Brassica juncea
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY: bjCBF1 gene
  - (B) LOCATION:

(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:  
 TTTCACCCTA TCTACCGGGG AGTTCGCCTG AGAAAGTCAG GTAAGTGGGT 50  
 GTGTGAAGTG AGGGAGCCAA ACAAGAAATC TAGGATTTGG CTTGGAACCT 100  
 TCAAAACCGC AGAGATCGCT GCTCGTGCTC ACGACGTTGC CGCCTTAGCC 150  
 CTCCGTGGAA GAGCGGCCTG TCTCAACTTC GCCGACTCGG CTTGGCGGCT 200  
 CCGTATCCCG GAGACAACTT GCGCCAAGGA TATCCAGAAG GCTGCTGCTG 250  
 AAGCTGCGTT GGCTTTTGGG GCCGAAAAGA GTGATACCAC GACGAATGAT 300  
 CAAGGCATGA ACATGGAGGA GATGACGGTG GTGGCTTCTC AGGCTGAGGT 350  
 GAGCGACACG ACGACATATC ATGGCCTGGA CATGGAGGAG ACTATGGTGG 400  
 AGGCTGTTTT TGCTGAGGAA CAGAGAGAAG GGTTTTACTT GGCGGAGGAG 450  
 ACGACGGTGG AGGGTGTTGT TACGGAGGAA CAGAGCAAAG GGTTTTATAT 500  
 GTACGAGGAG TGGACGTTCT GGATGCAGTC CTTTTTGGCC GATATGGCTG 550  
 AAGGCATGCT CTTTTCAAAG GGCGAAT 577

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CATCCGATCT ACAGGGGAGT TCGTCTGAGA AAATCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAACCAAACA AGAGATCTAG GATCTGGCTC GGTACTTTCC	100
TAACCGCCGA GATCGCAGCT CGCGCTCAGC ACGTCGCCGC CATAGCCCTC	150
CGTGGCAAAT CCGCATGTCT CAATTTGCT GACTCGGCTT GCGGGCTCCG	200
TATCTCGGAG ACAACATGCC CTAAGGAGAT TCAGAAGGCT GCTGCTGAAG	250
CCGCGGTGGC TTTTCAGGCT GAGCTAAATG ATACGACGGC CGATCATGGC	300
CTTGACGTGG AGGAGACGAT CGTGGAGGCT ATTTTCACGG AGGAAAGCAG	350
CGAAGGGTTT TATATGGACG AGGAGTTCAT GTTCGGGATG CCGACCTTGT	400
GGGCTAGTAT GGCAGAAGGG ATGCTTCTTC C	431

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: BJCBF2-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	Trp	
				5					10					15	
Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Arg	Ser	Arg	Ile	Trp	Leu	
				20					25					30	
Gly	Thr	Phe	Leu	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	Val	
				35					40					45	
Ala	Ala	Ile	Ala	Leu	Arg	Gly	Lys	Ser	Ala	Cys	Leu	Asn	Phe	Ala	
				50					55					60	
Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Ser	Glu	Thr	Thr	Cys	Pro	Lys	
				65					70					75	
Glu	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Val	Ala	Phe	Gln	Ala	
				80					85					90	
Glu	Leu	Asn	Asp	Thr	Thr	Ala	Asp	His	Gly	Leu	Asp	Val	Glu	Glu	
				95					100					105	
Thr	Ile	Val	Glu	Ala	Ile	Phe	Thr	Glu	Glu	Ser	Ser	Glu	Gly	Phe	
				110					115					120	
Tyr	Met	Asp	Glu	Glu	Phe	Met	Phe	Gly	Met	Pro	Thr	Leu	Trp	Ala	
				125					130					135	
Ser	Met	Ala	Glu	Gly	Met	Leu	Leu								
				140					145						

## (2) INFORMATION FOR SEQ ID NO:42:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 431
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bjCBF3 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CATCCAATTT ACCGTGGAGT TCGTCTGAGA AAATCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAGCCAAACA AGAAATCTAG GATCTGGCCC GGTACTTTCC	100
TAACCGCCGA GATCGCAGCT CGCGCTCAGC ACGTCGCCGC CATAGCCCTC	150
CGTGGCAAAT CCGCATGTCT CAATTTTCGCT GACTCGGCTT GGCGGCTCCG	200
TATCCCGGAG ACAACATGCC CTAAGGAGAT TCAGAAGGCT GCTGCTGAAG	250
CCGCGGTGGC TTTTCAGGCT GAGCTAAATG ATACGACGGC CGATCATGGC	300
CTTGACGTGG AGGAGACGAT CGTGGAGGCT ATTTTCACGG AGGAAAGCAG	350
CGAAGGGTTT TATATGGACG AGGAGTTCAT GTTCGGGATG CCGACCTTGT	400
GGGCTAGTAT GGCGGAGGGC ATGCTCCTTC C	431

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: BJCBF3-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	Trp
					5					10				15

Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	Pro	
				20					25					30	
Gly	Thr	Phe	Leu	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	Val	
				35					40					45	
Ala	Ala	Ile	Ala	Leu	Arg	Gly	Lys	Ser	Ala	Cys	Leu	Asn	Phe	Ala	
				50					55					60	
Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	Pro	Lys	
				65					70					75	
Glu	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Val	Ala	Phe	Gln	Ala	
				80					85					90	
Glu	Leu	Asn	Asp	Thr	Thr	Ala	Asp	His	Gly	Leu	Asp	Val	Glu	Glu	
				95					100					105	
Thr	Ile	Val	Glu	Ala	Ile	Phe	Thr	Glu	Glu	Ser	Ser	Glu	Gly	Phe	
				110					115					120	
Tyr	Met	Ala	Glu	Glu	Phe	Met	Phe	Gly	Met	Pro	Thr	Leu	Trp	Ala	
				125					130					135	
Ser	Val	Ala	Glu	Gly	Met	Leu	Leu								
				140					145						

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 425
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bxCBF4 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CATCCAATCT	ACCGGGGTGT	TCGACAGAGA	AACTCAGGGA	AATGGGTTTG	50
TGAAGTTAGG	GAGCCTAATA	AGAAATCTAG	GATCTGGTTA	GGGACTTTTC	100

```

CGACCGTCGA AATGGCCGCT CGTGCTCACG ACGTCGCCGC TTTAGCCCTT      150
CGTGGCCGCT CCGCTTGTCT TAATTTGCGC GACTCGGCGT GGTGTCTACG      200
GATTCCCGAG TCTACTTGTC CTAAAGAGAT TCAGAAAGCT GCGGCCGAAG      250
CTGCAATGGC GTTTCAGAAC GAGACGGCTA CGACTGAGAC GACTATGGTT      300
GAGGGAGTCA TACCGGCGGA GGAGACGGTG GGGCAGACGC GTGTGGAGAC      350
AGCAGAGGAG AACGGTGTGT TTTATATGGA CGATCCAAGG TTTCTTGAGA      400
ATATGGCAGA GGGCATGTTC CTACC                                     425

```

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 142
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica juncea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BJCBF4-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```

His Pro Ile Tyr Arg Gly Val Arg Gln Arg Asn Ser Gly Lys Trp
      5                      10                      15
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
      20                      25                      30
Gly Thr Phe Pro Thr Val Glu Met Ala Ala Arg Ala His Asp Val
      35                      40                      45
Ala Ala Leu Ala Leu Arg Gly Arg Ser Ala Cys Leu Asn Phe Ala
      50                      55                      60

```



Asp	Ser	Ala	Trp	Cys	Leu	Arg	Ile	Pro	Glu	Ser	Thr	Cys	Pro	Lys
				65					70					75
Glu	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Met	Ala	Phe	Gln	Asn
				80					85					90
Glu	Glu	Thr	Ala	Thr	Thr	Glu	Thr	Thr	Met	Val	Glu	Gly	Val	Ile
				95					100					105
Pro	Ala	Glu	Glu	Thr	Val	Gly	Gln	Thr	Arg	Val	Glu	Thr	Ala	Glu
				110					115					120
Glu	Asn	Gly	Val	Glu	Tyr	Met	Asp	Asp	Pro	Arg	Phe	Leu	Glu	Asn
				125					130					135
Met	Ala	Glu	Gly	Met	Leu	Phe								
				140					145					

## (2) INFORMATION FOR SEQ ID NO:46:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 632
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bnCBF1 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CACCCGATAT ACCGGGGAGT TCGTCTGAGA AAGTCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAACCAAACA AGAAATCTAG AATTTGGCTT GGAACCTTCA	100
AAACAGCTGA GATGGCAGCT CGTGCTCACG ACGTCGCTGC CCTAGCCCTC	150
CGTGGAAGAG GCGCCTGCCT CAATTATGCG GACTCGGCTT GGCGGCTCCG	200
CATCCCGGAG ACAACCTGCC ACAAGGATAT CCAGAAGGCT GCTGCTGAAG	250

```

CCGCATTGGC TTTTGAGGCT GAGAAAAGTG ATGTGACGAT GCAAATGGC      300
CAGAACATGG AGGAGACGAC GGCAGTGGCT TCTCAGGCTG AAGTGAATGA      350
CACGACGACA GAACATGGCA TGAACATGGA GGAGGCAACG GCAGTGGCTT      400
CTCAGGCTGA GGTGAATGAC ACGACGACGG ATCATGGCGT AGACATGGAG      450
GAGACAATGG TGGAGGCTGT TTTTACTGGG GAACAAAGTG AAGGGTTTAA      500
CATGGCGAAG GAGTCGACGG TGGAGGCTGC TGTGTGTACG GAGGAACCGA      550
GCAAAGGATC TTACATGGAC GAGGAGTGGG TGCTCGAGAT GCCGACCTTG      600
TTGGCTGATA TGGCAGAAGG GATGCTCCTG CC                          632

```

## (2) INFORMATION FOR SEQ ID NO:47:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BNCBF1-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```

His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
      5              10              15
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
      20              25              30
Gly Thr Phe Lys Thr Ala Glu Met Ala Ala Arg Ala His Asp Val
      35              40              45

```

Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	Leu	Asn	Tyr	Ala	
				55					55						60
Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	His	Lys	
				65					70						75
Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Glu	Ala	
				80					85						90
Glu	Lys	Ser	Asp	Val	Thr	Met	Gln	Asn	Gly	Gln	Asn	Met	Glu	Glu	
				95					100						105
Thr	Thr	Ala	Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	
				110					115						120
Glu	His	Gly	Met	Asn	Met	Glu	Glu	Ala	Thr	Ala	Val	Ala	Ser	Gln	
				125					130						135
Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Val	Asp	Met	Glu	
				140					145						150
Glu	Thr	Met	Val	Glu	Ala	Val	Phe	Thr	Gly	Glu	Gln	Ser	Glu	Gly	
				155					160						165
Phe	Asn	Met	Ala	Lys	Glu	Ser	Thr	Val	Glu	Ala	Ala	Val	Val	Thr	
				170					175						180
Glu	Glu	Pro	Ser	Lys	Gly	Ser	Tyr	Met	Asp	Glu	Glu	Trp	Met	Leu	
				185					190						195
Glu	Met	Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Glu	Gly	Met	Leu	Leu	
				200					205						210

## (2) INFORMATION FOR SEQ ID NO:48:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 876
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bnCBF2 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ACCGCTCGAG CAACAATGAA CACATTCCCT GCTTCCACTG AAATGGTTGG	50
CTCCGAGAAC GAGTCTCCGG TTACTACGGT AGTAGGAGGT GATTATTATC	100
CCATGTTGGC GGCAAGCTGT CCGAAGAAGC CAGCGGGTAG GAAGAAGTTT	150
CAGGAGACAC GTCACCCCAT TTACCGAGGA GTTCGTCTGA GAAAGTCAGG	200
TAAGTGGGTG TGTGAAGTGA GGGAACCAAA CAAGAAATCT AGAATTTGGC	250
CCGGAACTTT CAAAACAGCT GAGATGGCAG CTCGTGCTCA CGACGTCGCT	300
GCCCTAGCCC TCCGTGGAAG AGGCGCCTGC CTCAATTATG CGGACTCGGC	350
TTGGCGGCTC CGCATCCCGG AAACAACCTG CCACAAGGAT ATCCAGAAGG	400
CTGCTGCTGA AGCCGCATTG GCTTTTGAGG CTGAGAAAAG TGATGTGACG	450
ATGCAAAAATG GCCTGAACAT GGAGGAGACG ACGGCGGTGG CTTCTCAGGC	500
TGAAGTGAAT GACACGACGA CAGAACATGG CATGAACATG GAGGAGGCAA	550
CAGCGGTGGC TTCTCAGGCT GAGGTGAATG ACACGACGAC AGATCATGGC	600
GTAGACATGG AGGAGACGAT GGTGGAGGCT GTTTTTACGG AGGAACAAAG	650
TGAAGGGTTC AACATGGCGG AGGAGTCGAC GGTGGAGGCT GCTGTTGTTA	700
CGGATGAACT GAGCAAAGGA TTTTACATGG ACGAGGAGTG GACGTACGAG	750
ATGCCGACCT TGTTGGCTGA TATGGCGGCA GGGATGCTTT TGCCGCCACC	800
ATCTGTACAA TGGGGACATA ATGATGACTT GGAAGGAGAT GCGGACATGA	850
ACCTCTGGAG TTATTAAGGA TCCGCG	876

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: BNCBF2-PEP

(B) LOCATION:

(C) IDENTIFICATION METHOD: sequencing

(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Gly	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Tyr	Pro	Met	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Pro	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	His	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	Val	Thr	Met	Gln	Asn	Gly	Leu	140	145	150
Asn	Met	Glu	Glu	Thr	Ala	Val	Ala	Ser	Gln	Ala	Glu	Val	Asn		155	160	165
Asp	Thr	Thr	Thr	Glu	His	Gly	Met	Asn	Met	Glu	Glu	Ala	Thr	Ala	170	175	180
Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	185	190	195
Val	Asp	Met	Glu	Glu	Thr	Met	Val	Glu	Ala	Val	Phe	Thr	Glu	Glu	200	205	210
Gln	Ser	Glu	Gly	Phe	Asn	Met	Ala	Glu	Glu	Ser	Thr	Val	Glu	Ala	215	220	225
Ala	Val	Val	Thr	Asp	Glu	Leu	Ser	Lys	Gly	Phe	Tyr	Met	Asp	Glu	230	235	240
Glu	Trp	Thr	Tyr	Glu	Met	Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Ala	245	250	255
Gly	Met	Leu	Leu	Pro	Pro	Pro	Ser	Val	Gln	Trp	Gly	His	Asn	Asp	260	265	270
Asp	Leu	Glu	Gly	Asp	Ala	Asp	Met	Asn	Leu	Trp	Ser	Tyr			275	280	285

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 884

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Brassica napus
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY: bnCBF3 gene
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

ACTACACTCA GCCTTATCCA GTTTTTTTTCA AAAGATTTTT CAACAATGAA	50
CACATTCCCT GCGTCCACTG AAATGGTTGG CTCCGAGAAC GAGTCTCCGG	100
TTACTACGGT AGCAGGAGGT GATTATTATC CCATGTTGGC GGCAAGCTGT	150
CCGAAGAAGC CAGCAGGTAG GAAGAAGTTT CAGGAGACAC GTCACCCCAT	200
TTACCGAGGA GTTCGTCTGA GAAAGTCAGG TAAGTGGGTG TGTGAAGTGA	250
GGGAACCAAA CAAGAAATCT AGAATTGGC CCGGAAC TTT CAAAACAGCT	300
GAGATGGCAG CTCGTGCTCA CGACGTCGCT GCCCTAGCCC TCCGTGGAAG	350
AGGCGCCTGC CTCAATTATG CGGACTCGGC TTGGCGGCTC CGCATCCCGG	400
AGACAACCTG CCACAAGGAT ATCCAGAAGG CTGCTGCTGA AGCCGCATTG	450
GCTTTTGAGG CTGAGAAAAG TGATGTGACG ATGCAAAATG GCCTGAACAT	500
GGAGGAGACG ACGGCGGTGG CTTCTCAGGC TGAAGTGAAT GACACGACGA	550
CAGAACATGG CATGAACATG GAGGAGGCAA CGGCAGTGGC TTCTCAGGCT	600
GAGGTGAATG ACACGACGAC GGATCATGGC GTAGACATGG AGGAGACAAT	650
GGTGGAGGCT GTTTTTACTG GGAACAAAG TGAAGGGTTT AACATGGCGA	700
AGGAGTCGAC GGTGGAGGCT GCTGTTGTTA CGGAGGAACC GAGCAAAGGA	750

TCTTACATGG ACGAGGAGTG GATGCTCGAG ATGCCGACCT TGTTGGCTGA 800  
 TATGGCGGAA GGGATGCTTT TGCCGCCGCC GTCCGTACAA TGGGGACAGA 850  
 ATGATGACTT CGAAGGAGAT GCTGACATGA ACCT 884

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BNCBF3-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Gly	Ser	Glu	Asn
				5					10					15
Glu	Ser	Pro	Val	Thr	Thr	Val	Ala	Gly	Gly	Asp	Tyr	Tyr	Pro	Met
				20					25					30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe
				35					40					45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys
				50					55					60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser
				65					70					75
Arg	Ile	Trp	Pro	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg
				80					85					90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys
				95					100					105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr

Thr Cys His Lys	110	115	120
Asp Ile Gln Lys Ala	125	130	135
Ala Phe Glu Ala Glu Lys Ser Asp Val	140	145	150
Asn Met Glu Glu Thr Thr Ala Val Ala	155	160	165
Asp Thr Thr Thr Glu His Gly Met Asn	170	175	180
Val Ala Ser Gln Ala Glu Val Asn Asp	185	190	195
Val Asp Met Glu Glu Thr Met Val Glu	200	205	210
Gln Ser Glu Gly Phe Asn Met Ala Lys	215	220	225
Ala Val Val Thr Glu Glu Pro Ser Lys	230	235	240
Glu Trp Met Leu Glu Met Pro Thr Leu	245	250	255
Gly Met Leu Leu Pro Pro Pro Ser Val	260	265	270
Asp Phe Glu Gly Asp Ala Asp Met Asn	275	280	

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 874
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bnCBF4 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTAATTCGAT TACCGCTCGA GTACTACTA TACTACACTC AGCCTTATCC	50
AGTTTTTCAA AAGAAGTTTT CAACTATGAA CTCAGTCTCT ACTTTTTCTG	100
AACTTCTTGG CTCTGAGAAC GAGTCTCCGG TAGGTGGTGA TTACTGTCCC	150
ATGTTGGCGG CGAGCTGTCC GAAGAAGCCG GCGGGTAGGA AGAAGTTTCG	200
GGAGACACGT CACCCCATTT ACCGAGGAGT TCGCCTTAGA AAATCAGGTA	250
AGTGGGTGTG TGAAGTGAGG GAACCAAACA AAAAATCTAG GATTTGGCTC	300
GGAACTTTCA AAACAGCTGA GATCGCAGCT CGTGCTCACG ACGTCGCCGC	350
CTTAGCTCTC CGTGGAAGAG GCGCCTGCCT CAACTTCGCC GACTCGGCTT	400
GGCGGCTCCG TATCCCGGAG ACAACCTGCG CCAAGGATAT CCAGAAGGCT	450
GCTGCTGAAG CCGCATTGGC TTTTGAGGCC GAGAAGAGTG ATACCACGAC	500
GAATGATCAT GGCATGAACA TGGCTTCTCA GGCCGAGGTT AATGACACAA	550
CGGATCATGG CCTGGACATG GAGGAGACGA TGGTGGAGGC TGTTTTTACT	600
GAGGAGCAGA GAGACGGGTT TTACATGGCG GAGGAGACGA CGGTGGAGGG	650
TGTTGTTCCG GAGGAACAGA TGAGCAAAGG GTTTTACATG GACGAGGAGT	700
GGATGTTCCG GATGCCGACC TTGTTGGCTG ATATGGCGGC AGGGATGCTC	750
TTACCGCCGC CGTCCGTACA ATGGGGACAT AATGATGACT TCGAAGGAGA	800
TGTTGACATG AACCTCTGGA ATTATTAGTA CTCATATTTT TTAAATTAT	850
TTTTTGAACG AATAATATTT TATT	874

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A

(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: BNCBF4-PEP

(B) LOCATION:

(C) IDENTIFICATION METHOD: sequencing

(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met	Asn	Ser	Val	Ser	Thr	Phe	Ser	Glu	Leu	Leu	Gly	Ser	Glu	Asn	
				5					10					15	
Glu	Ser	Pro	Val	Gly	Gly	Asp	Tyr	Cys	Pro	Met	Leu	Ala	Ala	Ser	
				20					25					30	
Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	Thr	Arg	
				35					40					45	
His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	Trp	
				50					55					60	
Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	Leu	
				65					70					75	
Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	Val	
				80					85					90	
Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	Leu	Asn	Phe	Ala	
				95					100					105	
Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	Ala	Lys	
				110					115					120	
Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Glu	Ala	
				125					130					135	
Glu	Lys	Ser	Asp	Thr	Thr	Thr	Asn	Asp	His	Gly	Met	Asn	Met	Ala	
				140					145					150	
Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Asp	His	Gly	Leu	Asp	Met	
				155					160					165	
Glu	Glu	Thr	Met	Val	Glu	Ala	Val	Phe	Thr	Glu	Glu	Gln	Arg	Asp	
				170					175					180	
Gly	Phe	Tyr	Met	Ala	Glu	Glu	Thr	Thr	Val	Glu	Gly	Val	Val	Pro	
				185					190					195	
Glu	Glu	Gln	Met	Ser	Lys	Gly	Phe	Tyr	Met	Asp	Glu	Glu	Trp	Met	
				200					205					210	
Phe	Gly	Met	Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Ala	Gly	Met	Leu	
				215					220					225	
Leu	Pro	Pro	Pro	Ser	Val	Gln	Trp	Gly	His	Asn	Asp	Asp	Phe	Glu	
				230					235					240	
Gly	Asp	Val	Asp	Met	Asn	Leu	Trp	Asn	Tyr						
				245					250						

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 898

(B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: bnCBF5 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AATAAATATC TTATCAAACC AGTCAGAACA GAGATCTTGT TACTTACTAT	50
ACTACACTCA GCCTTATCCA GTTTTCAAAA AAAGTATTCA ACGATGAACT	100
CAGTCTCTAC TTTTCTGAA CTGCTCCGCT CCGAGAACGA GTCTCCGGTT	150
AATACGGAAG GTGGTGATTA CATTTTGGCG GCGAGCTGTC CCAAGAAACC	200
TGCTGGTAGG AAGAAGTTTC AGGAGACACG CCACCCCAT TACAGAGGAG	250
TTCGTCTGAG GAAGTCAGGT AAGTGGGTGT GTGAAGTGAG GGAACCAAAC	300
AAGAAATCTA GAATTTGGCT CGGAAC TTTC AAAACAGCTG AGATCGCAGC	350
TCGTGCTCAC GACGTTGCCG CCTTAGCTCT CCGTGGAAGA GGCGCCTGCC	400
TCAACTTCGC CGACTCGGCT TGGCGGCTCC GTATCCCGGA GACGACCTGC	450
GCCAAGGATA TCCAGAAGGC TGCTGCTGAA GCCGCATTGG CTTTGTAGGC	500
CGAGAAGAGT GATACCACGA CGAATGATCA TGGCATGAAC ATGGCTTCTC	550
AGGTTGAGGT TAATGACACG ACGGATCATG ACCTGGACAT GGAGGAGACG	600
ATAGTGGAGG CTGTTTTTAG GGAGGAACAG AGAGAAGGGT TTTACATGGC	650
GGAGGAGACG ACGGTTGTGG GTGTTGTTCC GGAGGAACAG ATGAGCAAAG	700
GGTTTTACAT GGACGAGGAG TGGATGTTCC GGATGCCGAC CTTGTTGGCT	750
GATATGGCGG CAGGGATGCT CTTACCGCTG CCGTCCGTAC AATGGGGACA	800
TAATGATGAC TTCGAAGGAG ATGCTGACAT GAACCTCTGG AATTATTAGT	850

ACTCATATTT TTTTAAATTA TTTTTTGAAC GAATAATATT TTATTGAA

898

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 251
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: BNCBF5-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Met	Asn	Ser	Val	Ser	Thr	Phe	Ser	Glu	Leu	Leu	Arg	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Asn	Thr	Glu	Gly	Gly	Asp	Tyr	Ile	Leu	Ala	Ala	20	25	30
Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Gln	Glu	Thr	35	40	45
Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	50	55	60
Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	65	70	75
Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	80	85	90
Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	Leu	Asn	Phe	95	100	105
Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	Ala	110	115	120
Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Glu	125	130	135
Ala	Glu	Lys	Ser	Asp	Thr	Thr	Thr	Asn	Asp	His	Gly	Met	Asn	Met			

Ala	Ser	Gln	Val	140	Glu	Val	Asn	Asp	Thr	145	Thr	Asp	His	Asp	Leu	Asp	150
Met	Glu	Glu	Thr	155	Ile	Val	Glu	Ala	Val	160	Phe	Arg	Glu	Glu	Gln	Arg	165
Glu	Gly	Phe	Tyr	170	Met	Ala	Glu	Glu	Thr	175	Thr	Val	Val	Gly	Val	Val	180
Pro	Glu	Glu	Gln	185	Met	Ser	Lys	Gly	Phe	190	Tyr	Met	Asp	Glu	Glu	Trp	195
Met	Phe	Gly	Met	200	Pro	Thr	Leu	Leu	Ala	205	Asp	Met	Ala	Ala	Gly	Met	210
Leu	Leu	Pro	Leu	215	Pro	Ser	Val	Gln	Trp	220	Gly	His	Asn	Asp	Asp	Phe	225
Glu	Gly	Asp	Ala	230	Asp	Met	Asn	Leu	Trp	235	Asn	Tyr					240
				245						250							255

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1132
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bnCBF6 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GATTACCGCT	CGAGTACTTA	CTATACTACA	CTCAGCCTTA	TCCAGTTTTT	50
CTCAAAAGAT	TTTTCAACAA	TGAACACATT	CCCTGCTTCC	ACTGAAATGG	100
TTGGCTCCGA	GAACGAGTCT	CCGGTTACTA	CGGTAGTAGG	AGGTGATTAT	150

TATCCCATGT TGGCGGCAAG CTGTCCGAAG AAGCCAGCGG GTAGGAAGAA	200
GTTTCAGGAG ACACGTCACC CCATTTACCG AGGAGTTCGT CTGAGAAAGT	250
CAGGTAAGTG GGTGTGTGAA GTGAGGGAAC CAAACAAGAA ATCTAGAATT	300
TGGCTTGGAA CTTTCAAAAC AGCTGAGATG GCAGCTCGTG CTCACGACGT	350
GGCTGCCCTA GCCCTCCGTG GAAGAGGCGC CTGCCTCAAT TATGCGGACT	400
CGGCTTCGCG GCTCCGCATC CCGGAGACAA CCTGCCACAA GGATATCCAG	450
AAGGCTGCTG CTGAAGCCGC ATTGGCTTTT GAGGCTGAGA AAAGTGATGT	500
GACGATGGAG GAGACGATGG CGGTGGCTTC TCAGGCTGAA GTGAATGACA	550
CGACGACAGA TCATGGCATG AACATGGAGG AGGCAACAGC GGTGGCTTCT	600
CAGGCTGAGG TGAATGACAC GACGACAGAT CATGGCGTAG ACATGGAGGA	650
GACGATGGTG GAGGCTGTTT TTACGGAGGA ACAAAGTGAA GGGTTCAACA	700
TGGCGGAGGA GTCGACGGTG GAGGCTGCTG TTGTTACGGA TGAAGTGAAG	750
AAAGGATTTT ACATGGACGA GGAGTGGACG TACGAGATGC CGACCTTGTT	800
GGCTGATATG GCGGCAGGGA TGCTTTTGCC GCCACCATCT GTACAATGGG	850
GACATAATGA TGAAGTGGAA GGAGATGCTG ACATGAACCT CTGGAATTAT	900
TAATACTCGT GTTTTAAAAA TTATACATTG TGCAATAATA TTTTATCGAA	950
TTTCTAATTC TGCCTTTAAC TTTTAATGGG GATCTTTATT AGTGTAGGAA	1000
ACGAGTGTAATGTTCCGCC GTGGTGTGTTGT CAAAATGCTG ATTATTTTTT	1050
GTGTGCAGCA TAATCACGTT TGGTTTCCTT TACACTCCAA ATTTAGTTGA	1100
AATACAAATA GAATAGAAAA GTGAAAAAAT GT	1132

## (2) INFORMATION FOR SEQ ID NO:57:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A

(F) TISSUE TYPE: N/A  
 (G) CELL TYPE: N/A  
 (H) CELL LINE: N/A  
 (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
 (A) NAME/KEY: BNCBF6-PEP  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Gly	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Tyr	Pro	Met	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Ser	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	His	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	Val	Thr	Met	Glu	Glu	Thr	Met	140	145	150
Ala	Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	155	160	165
Gly	Met	Asn	Met	Glu	Glu	Ala	Thr	Ala	Val	Ala	Ser	Gln	Ala	Glu	170	175	180
Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Val	Asp	Met	Glu	Glu	Thr	185	190	195
Met	Val	Glu	Ala	Val	Phe	Thr	Glu	Glu	Gln	Ser	Glu	Gly	Phe	Asn	200	205	210
Met	Ala	Glu	Glu	Ser	Thr	Val	Glu	Ala	Ala	Val	Val	Thr	Asp	Glu	215	220	225
Leu	Ser	Lys	Gly	Phe	Tyr	Met	Asp	Glu	Glu	Trp	Thr	Tyr	Glu	Met	230	235	240
Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Ala	Gly	Met	Leu	Leu	Pro	Pro	245	250	255
Pro	Ser	Val	Gln	Trp	Gly	His	Asn	Asp	Asp	Leu	Glu	Gly	Asp	Ala	260	265	270
Asp	Met	Asn	Leu	Trp	Asn	Tyr									275	280	

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 768  
 (B) TYPE: Nucleic Acid

(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Brassica napus*  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: bnCBF7 gene  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AGTGATGTTT TTCAAAAGAA GTTTTCAACT ATGAAGTCAG TCTCTACTTT	50
TTCTGAACTT CTTGGCTCTG AGAACGAGTC TCCGGTAGGT GGTGATTACT	100
GTCCCATGTT GCGGCGGAGC TGTCCGAAGA AGCCGGCGGG TAGGAAGAAG	150
TTTCGGGAGA CACGTCACCC CATTTACCGA GGAGTTCGCC TTAGAAAATC	250
GGCTCGGTAC TTTCCCTAACA GCCGAGATCG CAGCCCGTGC TCACGACGTC	300
GCCGCCATAG CCCTCCGCGG CAAATCAGCT TGTCTCAATT TTGCCGACTC	350
CGCTTGCGG CTCCGTATCC CGGAGACAAC ATGCCCAAG GAGATTCAGA	400
AGGCGGCTGC TGAAGCCGCG GTGGCTTTTA AGGCTGAGAT AAATAATACG	450
ACGGCGGATC ATGGCATTGA CGTGGAGGAG ACGATCGTTG AGGCTATTTT	500
CACGGAGGAA AACAACGATG GTTTTTATAT GGACGAGGAG GAGTCCATGT	550
TCGGGATGCC GGCCTTGTTG GCTAGTATGG CTGAAGGAAT GCTTTTGCCG	600
CCTCCGTCCG TACAATTCGG ACATACCTAT GACTTTGACG GAGATGCTGA	650
CGTGTCCTT TGGAGTTATT AGTACAAAGA TTTTTATTT CCATTTTTGG	700
TATAATACTT CTTTTTGATT TTCGGATTCT ACCTTTTTAT GGGTATCATT	750



TTTTTTT TAG GAAACGGG

768

## (2) INFORMATION FOR SEQ ID NO:59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BNCBF7-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

Met Asn Ser Val Ser Thr Phe Ser Glu Leu Leu Gly Ser Glu Asn
      5              10              15
Glu Ser Pro Val Gly Gly Asp Tyr Cys Pro Met Leu Ala Ala Ser
      20              25              30
Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu Thr Arg
      35              40              45
His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
      50              55              60
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
      65              70              75
Gly Thr Phe Leu Thr Ala Glu Ile Ala Ala Arg Ala His Asp Val
      80              85              90
Ala Ala Ile Ala Leu Arg Gly Lys Ser Ala Cys Leu Asn Phe Ala
      95              100             105
Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Pro Lys
      110             115             120
Glu Ile Gln Lys Ala Ala Ala Glu Ala Val Ala Phe Lys Ala
      125             130             135
Glu Ile Asn Asn Thr Thr Ala Asp His Gly Ile Asp Val Glu Glu

```

55/101

Thr	Ile	Val	Glu	140	Ala	Ile	Phe	Thr	Glu	145	Glu	Asn	Asn	Asp	Gly	150	Phe
				155						160							
Tyr	Met	Asp	Glu	170	Glu	Glu	Ser	Met	Phe	175	Gly	Met	Pro	Ala	Leu	180	Leu
Ala	Ser	Met	Ala	185	Glu	Gly	Met	Leu	Leu	190	Pro	Pro	Pro	Ser	Val	195	Gln
Phe	Gly	His	Thr	200	Tyr	Asp	Phe	Asp	Gly	205	Asp	Ala	Asp	Val	Ser	210	Leu
Trp	Ser	Tyr															

215

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 953
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: bnCBF8 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ACCGCTCGAG CAACAATGAA CACATTCCCT GCTTCCACTG AAATGGTTGG	50
CTCCGAGAAC GAGTCTCCGG TTACTACGGT AGCAGGAGGT GATTATTATC	100
CCATGTTGGC GGCAAGCTGT CCGAAGAAGC CAGCGGGTAG GAAGAAGTTT	150
CAGGAGACAC GTCACCCCAT TTACCGAGGA GTTCGTCTGA GAAAGTCAGG	200
TAAGTGGGTG TGTGAAGTGA GGGAACCAAA CAAGAAATCT AGAATTTGGC	250

TTGGAAC TTT CAAAACAGCT GAGATGGCAG CTCGTGCTCA CGACGTGGCT	300
GCCCTAGCCC TCCGTGGAAG AGGCGCCTGC CTCAATTATG CGGACTCGGC	350
TTCGCGGCTC CGCATCCCGG AGACAACCTG CCACAAGGAT ATCCAGAAGG	400
CTGCTGCTGA AGCCGCATTG GCTTTTGAGG CTGAGAAAAG TGATGTGACG	450
ATGGAGGAGA CGATGGCGGT GGCTTCTCAG GCTGAAGTGA ATGACACGAC	500
GACAGATCAT GGCATGAACA TGGAGGAGGC AACGGCAGTG GCTTCTCAGG	550
CTGAGGTGAA TGACACGACG ACGGATCATG GCGTAGACAT GGAGGAGACA	600
ATGGTGGAGG CTGTTTTTAC TGGGGAACAA AGTGAAGGGT TTAACATGGC	650
GAAGGAGTCG ACGGTGGAGG CTGCTGTTGT TACGGAGGAA CCGAGCAAAG	700
GATCTTACAT GGACGAGGAG TGGATGCTCG AGATGCCGAC CTTGTTGGCT	750
GATATGGCGG AAGGGATGCT TTTGCCGCCG CCGTCCGTAC AATGGGGACA	800
GAATGATGAC TTCGAAGGAG ATGCGGACAT GAACCTCTGG AGTTATTAAT	850
ACTCGTATTT TTAAAATTAT TTATTGTGCA ATAATTTTTT ATCGAATTTT	900
GAATTCTGCC TTTAATTTTT AATGGGGATC TTTATTTGCC AAAAAAAAAA	950
AAA	953

## (2) INFORMATION FOR SEQ ID NO:61:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

(A) NAME/KEY: BNCBF8-PEP  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Gly	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Ala	Gly	Gly	Asp	Tyr	Tyr	Pro	Met	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Ser	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	His	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	Val	Thr	Met	Glu	Glu	Thr	Met	140	145	150
Ala	Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	155	160	165
Gly	Met	Asn	Met	Glu	Glu	Ala	Thr	Ala	Val	Ala	Ser	Gln	Ala	Glu	170	175	180
Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Val	Asp	Met	Glu	Glu	Thr	185	190	195
Met	Val	Glu	Ala	Val	Phe	Thr	Gly	Glu	Gln	Ser	Glu	Gly	Phe	Asn	200	205	210
Met	Ala	Lys	Glu	Ser	Thr	Val	Glu	Ala	Ala	Val	Val	Thr	Glu	Glu	215	220	225
Pro	Ser	Lys	Gly	Ser	Tyr	Met	Asp	Glu	Glu	Trp	Met	Leu	Glu	Met	230	235	240
Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Pro	245	250	255
Pro	Ser	Val	Gln	Trp	Gly	Gln	Asn	Asp	Asp	Phe	Glu	Gly	Asp	Ala	260	265	270
Asp	Met	Asn	Leu	Trp	Ser	Tyr									275	280	

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 889  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Brassica napus*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: bnCBF9 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CTAGTGATTA CCGCTCGAGC AACAAATGAAC ACATTCCTG CTTCCACTGA	50
AATGGTTGGC TCCGAGAACG AGTCTCCGGT TACTACGGTA GCAGGAGGTG	100
ATTATTATCC CATGTTGGCG GCAAGCTGTC CGAAGAAGCC AGCGGGTAGG	150
AAGAAGTTTC AGGAGACACG TCACCCCATTT TACCGAGGAG TTCGTCTGAG	200
AAAGTCAGGT AAGTGGGTGT GTGAAGTGAG GGAACCAAAC AAGAAATCTA	250
GAATTTGGCC CGGAACTTTC AAAACAGCTG AGATGGCAGC TCGTGCTCAC	300
GACGTCGCTG CCCTAGCCCT CCGTGGAAGA GGCGCCCGCC TCAATTATGC	350
GGACTCAGCT TGGCGGCTCC GCATCCCGGA GACAACCTGC CACAAGGATA	400
TCCAGAAGGC TGCTGCTGAA GCCGCATTGG CTTTGTAGGC TGAGAAAAGT	450
GATGTGACGA TGCAAAATGG CCTGAACATG GAGGAGACGA CGGCGGTGGC	500
TTCTCAGGCT GAAGTGAATG ACACGACGAC AGAACATGGC ATGAACATGG	550
AGGAGGCAAC GGCAGTGGCT TCTCAGGCTG AGGTGAATGA CACGACGACG	600
GATCATGGCG TAGACATGGA GGAGACAATG GTGGAGGCTG TTTTACTGG	650
GGAACAAAGT GAAGGGTTTA ACATGGCGAA GGAGTCGACG GTGGAGGCTG	700
CTGTTGTTAC GGAGGAACCG AGCAAAGGAT CTTACATGGA CGAGGAGTGG	750
ATGCTCGAGA TGCCGACCTT GTTGGCTGAT ATGGCGGAAG GGATGCTTTT	800
GCCGCCGCCG TCCGTACAAT GGGGACAGAA TGATGACTTC GAAGGAGATG	850
CGCACATGAA CCTCTGGAGT TATTAAGGAT CCGCGAATC	889

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 283
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica napus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BNCBF9-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```

Met Asn Thr Phe Pro Ala Ser Thr Glu Met Val Gly Ser Glu Asn
      5      10
Glu Ser Pro Val Thr Thr Val Ala Gly Asp Tyr Tyr Pro Met
      20      25
Leu Ala Ala Ser Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe
      35      40
Gln Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys
      50      55
Ser Gly Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser
      65      70
Arg Ile Trp Pro Gly Thr Phe Lys Thr Ala Glu Met Ala Ala Arg
      80      85
Ala His Asp Val Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Arg
      95     100
Leu Asn Tyr Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr
     110     115
Thr Cys His Lys Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu
     125     130
Ala Phe Glu Ala Glu Lys Ser Asp Val Thr Met Gln Asn Gly Leu
     140     145
Asn Met Glu Glu Thr Thr Ala Val Ala Ser Gln Ala Glu Val Asn

```

	155		160		165
Asp Thr Thr Thr	Glu His Gly Met Asn Met	Glu Glu Ala Thr	Ala		
	170		175		180
Val Ala Ser Gln	Ala Glu Val Asn Asp Thr Thr	Thr Asp His	Gly		
	185		190		195
Val Asp Met Glu	Glu Thr Met Val Glu Ala	Val Phe Thr Gly	Glu		
	200		205		210
Gln Ser Glu Gly	Phe Asn Met Ala Lys Glu	Ser Thr Val Glu	Ala		
	215		220		225
Ala Val Val Thr	Glu Glu Pro Ser Lys Gly	Ser Tyr Met Asp	Glu		
	230		235		240
Glu Trp Met Leu	Glu Met Pro Thr Leu Leu	Ala Asp Met Ala	Glu		
	245		250		255
Gly Met Leu Leu	Pro Pro Pro Ser Val Gln	Trp Gly Gln Asn	Asp		
	260		265		270
Asp Phe Glu Gly	Asp Ala His Met Asn Leu	Trp Ser Tyr			
	275		280		285

## (2) INFORMATION FOR SEQ ID NO:64:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 563
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: boCBF1 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CACCCTATCT ACCGGGGAGT TCGCCTGAGA AAGTCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAGCCAAACA AGAAATCTAG GATTTGGCTT GGAAC TTCA	100

```

AAACCGCAGA GATCGCTGCT CGTGCTCACG ACGTTGCCGC CTTAGCCCTC      150
CGTGGAAGAG CGGCCTGTCT CAACTTCGCC GACTCGGCTT GCGGGCTCCG      200
TATCCCGGAG ACAACTTGCG CCAAGGATAT CCAGAAGGCT GCTGCTGAAG      250
CTGCGTTGGC TTTTGGGGCC GAAAAGAGTG ATACCACGAC GAATGATCAA      300
GGCATGAACA TGGAGGAGAT GACGGTGGTG GCTTCTCAGG CTGAGGTGAG      350
CGACACGACG ACATATCATG GCCTGGACAT GGAGGAGACT ATGGTGGAGG      400
CTGTTTTTGC TGAGGAACAG AGAGAAGGGT TTTACTTGGC GGAGGAGACG      450
ACGGTGGAGG GTGTTGTTAC GGAGGAACAG AGCAAAGGGT TTTATATGGA      500
CGAGGAGTGG ACGTTCGGGA TGCAGTCCTT TTTGGCCGAT ATGGCTGAAG      550
GCATGCTCTT TCC                                              563

```

## (2) INFORMATION FOR SEQ ID NO:65:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 188
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BOCBF1-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```

His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
          5                      10                      15

```



```

Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
      20      25      30
Gly Thr Phe Lys Thr Ala Glu Ile Ala Ala Arg Ala His Asp Val
      35      40      45
Ala Ala Leu Ala Leu Arg Gly Arg Ala Ala Cys Leu Asn Phe Ala
      50      55      60
Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Ala Lys
      65      70      75
Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Gly Ala
      80      85      90
Glu Lys Ser Asp Thr Thr Thr Asn Asp Gln Gly Met Asn Met Glu
      95     100     105
Glu Met Thr Val Val Ala Ser Gln Ala Glu Val Ser Asp Thr Thr
     110     115     120
Thr Tyr His Gly Leu Asp Met Glu Glu Thr Met Val Glu Ala Val
     125     130     135
Phe Ala Glu Glu Gln Arg Glu Gly Phe Tyr Leu Ala Glu Glu Thr
     140     145     150
Thr Val Glu Gly Val Val Thr Glu Glu Gln Ser Lys Gly Phe Tyr
     155     160     165
Met Asp Glu Glu Trp Thr Phe Gly Met Gln Ser Phe Leu Ala Asp
     170     175     180
Met Ala Glu Gly Met Leu Phe Pro
           185

```

## (2) INFORMATION FOR SEQ ID NO:66:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 533
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: boCBF2 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GAAACATAGA TCTTTGTACT TACTATACTT CACCTTATCC AGTTTTATTT	50
TTTTATTTAT AAAGAGTTTT CAACAATGAC CTCATTTTCT ACCTTTTCTG	100
AACTGTTGGG CTCCGAGCAT GAGTCTCCGG TTACATTAGG CGAAGAGTAT	150
TGTCCGAAGC TGGCCGCAAG CTGTCCGAAG AAACCAGCCG GCCGGAAGAA	200
GTTTCGAGAG ACGCGTCACC CAGTTTACAG AGGAGTTCGT CTGAGAAACT	250
CAGGTAAGTG GGTGTGTGAA GTGAGGGAGC CAAACAAGAA ATCTAGGATT	300
TGGCTCGGTA CTTTCCTAAC AGCCGAGATC GCAGCCCGTG CTCACGACGT	350
CGCCGCCATA GCCCTCCGCG GCAAATCAGC TTGTCTCAAT TTTGCCGACT	400
CCGCTTGGCG GCTCCGTATC CCGGAGACAA CATGCCCCAA GGAGATTGAG	450
AAGGCGGCTG CTGAAGCCGC GGTGGCTTTT AAGGCTGAGA TAAATAATAC	500
GACGGCGGAT CACGGCCTCG ACATGGAAGA GAC	533

## (2) INFORMATION FOR SEQ ID NO:67:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 152
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BOCBF2-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Met	Thr	Ser	Phe	Ser	Thr	Phe	Ser	Glu	Leu	Leu	Gly	Ser	Glu	His	
				5					10					15	
Glu	Ser	Pro	Val	Thr	Leu	Gly	Glu	Glu	Tyr	Cys	Pro	Lys	Leu	Ala	
				20					25					30	
Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	
				35					40					45	
Thr	Arg	His	Pro	Val	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Asn	Ser	Gly	
				50					55					60	
Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	
				65					70					75	
Trp	Leu	Gly	Thr	Phe	Leu	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	
				80					85					90	
Asp	Val	Ala	Ala	Ile	Ala	Leu	Arg	Gly	Lys	Ser	Ala	Cys	Leu	Asn	
				95					100					105	
Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	
				110					115					120	
Pro	Lys	Glu	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Val	Ala	Phe	
				125					130					135	
Lys	Ala	Glu	Ile	Asn	Asn	Thr	Thr	Ala	Asp	His	Gly	Leu	Asp	Met	
				140					145					150	
Glu	Glu														

155

## (2) INFORMATION FOR SEQ ID NO:68:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 887
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: boCBF3 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing

## (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

ACTCAGCCTT ATCCAGTTTT TCTCAAAGA TTTTCAACA ATGAACACAT	50
TCCCTGCTTC CACTGAAATG GTTGGCTCCG AGAACGAGTC TCCGTTACT	100
ACGGTAGTAG GAGGTGATTA TTATCCCATG TTGGCGGCAA GCTGTCCGAA	150
GAAGCCAGCG GGTAGGAAGA AGTTTCAGGA GACACGTCAC CCCATTTACC	200
GAGGAGTTCG TCTGAGAAAG TCAGGTAAGT GGGTGTGTGA AGTGAGGGAA	250
CCAAACAAGA AATCTAGAAT TTGGCTTGGA ACTTTCAAAA CAGCTGAGAT	300
GGCAGCTCGT GCTCACGACG TGGCTGCCCT AGCCCTCCGT GGAAGAGGCG	350
CCTGCCTCAA TTATGCGGAC TCGGCTTGGC GGCTCCGCAT CCCGGAGACA	400
ACCTGCCACA AGGATATCCA GAAGGCTGCT GCTGAAGCCG CATTGGCTTT	450
TGAGGCTGAG AAAAGTGATG TGACGATGGA GGAGACGATG GCGGTGGCTT	500
CTCAGGCTGA AGTGAATGAC ACGACGACAG ATCATGGCAT GAACATGGAG	550
GAGGCAACAG CGGTGGCTTC TCAGGCTGAG GTGAATGACA CGACGACAGA	600
TCATGGCGTA GACATGGAGG AGACGATGGT GGAGGCTGTT TTTACGGAGG	650
AACAAAGTGA AGGGTTCAAC ATGGCGGAGG AGTCGACGGT GGAGGCTGCT	700
GTTGTTACGG ATGAACTGAG CAAAGGATTT TACATGGACG AGGAGTGGAC	750
GTACGAGATG CCGACCTTGT TGGCTGATAT GGCAGCAGGG ATGCTTTTGC	800
CGCCACCATC TGTACAATGG GGACATAATG ATGACTTGGA AGGAGATGCG	850
GACATGAACC TCTGGAGTTA TTAATACTCG TATTTTTT	887

## (2) INFORMATION FOR SEQ ID NO:69:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A

(E) HAPLOTYPE: N/A  
 (F) TISSUE TYPE: N/A  
 (G) CELL TYPE: N/A  
 (H) CELL LINE: N/A  
 (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: BOCBF3-PEP  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Gly	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Tyr	Pro	Met	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	His	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	Val	Thr	Met	Glu	Glu	Thr	Met	140	145	150
Ala	Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	155	160	165
Gly	Met	Asn	Met	Glu	Glu	Ala	Thr	Ala	Val	Ala	Ser	Gln	Ala	Glu	170	175	180
Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Val	Asp	Met	Glu	Glu	Thr	185	190	195
Met	Val	Glu	Ala	Val	Phe	Thr	Glu	Glu	Gln	Ser	Glu	Gly	Phe	Asn	200	205	210
Met	Ala	Glu	Glu	Ser	Thr	Val	Glu	Ala	Ala	Val	Val	Thr	Asp	Glu	215	220	225
Leu	Ser	Lys	Gly	Phe	Tyr	Met	Asp	Glu	Glu	Trp	Thr	Tyr	Glu	Met	230	235	240
Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Ala	Gly	Met	Leu	Leu	Pro	Pro	245	250	255
Pro	Ser	Val	Gln	Trp	Gly	His	Asn	Asp	Asp	Leu	Glu	Gly	Asp	Ala	260	265	270
Asp	Met	Asn	Leu	Trp	Ser	Tyr									275	280	

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 950
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: boCBF4 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CTGAAAAGAA GATAAAAGAG AGAGAAATAA ATATCTTATC AAACCAGACA	50
GAACAGAGAT CTTGTTACTT ACTATACTAC ACTCAGCCTT ATCCAGTTTT	100
TCAAAAGAAG TTTTCAACTA TGAAGTCAGT CTCTACTTTT TCTGAACTTC	150
TTGGCTCTGA GAACGAGTCT CCGGTAGGTG GTGATTACTG TCCCATGTTG	200
GCGGCGAGCT GTCCGAAGAA GCCGGCGGGT AGGAAGAAGT TTCGGGAGAC	250
ACGTCACCCC ATTTACCGAG GAGTTCGCCT TAGAAAATCA GGTAAGTGGG	300
TGTGTGAAGT GAGGGAACCA AACAAAAAAT CTAGGATTTG GCTCGGAACT	350
TTCAAAACAG CTGAGATCGC AGCTCGTGCT CACGACGTCG CCGCCTTAGC	400
TCTCCGTGGA AGAGGCGCCT GCCTCAACTT CGCCGACTCG GCTTGGCGGC	450
TCCGTATCCC GGAGACAACC TGCGCCAAGG ATATCCAGAA GGCTGCTGCT	500
GAAGCCGCAT TGGCTTTTGA GGCCGAGAAG AGTGATACCA CGACGAATGA	550
TCATGGCATG AACATGGCTT CTCAGGCTGA GGTAAATGAC ACGACGGATC	600
ATGGCCTGGA CATGGAGGAG ACGATGGTGG AGGCTGTTTT TACTGAGGAG	650

```

CAGAGAGACG GGTTTTACAT GCGGAGGAG ACGACGGTGG AGGGTGTGTG      700
TCCGGAGGAA CAGATGAGCA AAGGGTTTTA CATGGACGAG GAGTGGATGT      750
TCGGGATGCC GACCTTGTTG GCTGATATGG CGGCAGGGAT GCTCTTACCG      800
CCGCCGTCCG TACAATGGGG ACATAATGAT GACTTCGAAG GAGATGCTGA      850
CATGAACCTC TGGAATTATT AGTACTCGTA TTTTTTTAAA TTATTTTTTG      900
AACGAATAAT ATTTTATTGA ATTCGGATTC TACCTGTTTT TTTAATGGAT      950

```

## (2) INFORMATION FOR SEQ ID NO:71:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BOCBF4-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```

Met Asn Ser Val Ser Thr Phe Ser Glu Leu Leu Gly Ser Glu Asn
      5              10              15
Glu Ser Pro Val Gly Gly Asp Tyr Cys Pro Met Leu Ala Ala Ser
      20              25              30
Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu Thr Arg
      35              40              45
His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
      50              55              60
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu

```

	65		70		75
Gly Thr Phe Lys Thr Ala Glu Ile Ala Ala Arg Ala His Asp Val					
	80		85		90
Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Cys Leu Asn Phe Ala					
	95		100		105
Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Ala Lys					
	110		115		120
Asp Ile Gln Lys Ala Ala Glu Ala Ala Leu Ala Phe Glu Ala					
	125		130		135
Glu Lys Ser Asp Thr Thr Thr Asn Asp His Gly Met Asn Met Ala					
	140		145		150
Ser Gln Ala Glu Val Asn Asp Thr Thr Asp His Gly Leu Asp Met					
	155		160		165
Glu Glu Thr Met Val Glu Ala Val Phe Thr Glu Glu Gln Arg Asp					
	170		175		180
Gly Phe Tyr Met Ala Glu Glu Thr Thr Val Glu Gly Val Val Pro					
	185		190		195
Glu Glu Gln Met Ser Lys Gly Phe Tyr Met Asp Glu Glu Trp Met					
	200		205		210
Phe Gly Met Pro Thr Leu Leu Ala Asp Met Ala Ala Gly Met Leu					
	215		220		225
Leu Pro Pro Pro Ser Val Gln Trp Gly His Asn Asp Asp Phe Glu					
	230		235		240
Gly Asp Ala Asp Met Asn Leu Trp Asn Tyr					
	245		250		

## (2) INFORMATION FOR SEQ ID NO:72:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 877
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: boCBF5 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:



## (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

ACCGCTCGAG CAACAATGAA CACATTCCCT GCTTCCACTG AAATGGTTAG	50
CTCCGAGAAC GAGTCTCCGG TTACTACGGT AGTAGGAGGT GATTATTATC	100
CCATGTTGGC GGCAAGCTGT CCGAAGAAGC CAGCGGGTAG GAAGAAGTTT	150
CAGGAGACAC GTCACCCCAT TTACCGAGGA GTTCGTCTGA GAAAGTCAGG	200
TAAGTGGGTG TGTGAAGTGA GGGAATAAA CAAGAAATCT AGAATTTGGC	250
TTGGAAC TTT CAAAACAGCT GAGATGGCAG CTCGTGCTCA CGACGTGGCT	300
GCCCTAGCCC TCCGTGGAAG AGGCGCCTGC CTCAATTATG CGGACTCGGC	350
TTGGCGGCTC CGCATCCCGG AGACAACCTG CCACAAGGAT ATCCAGAAGG	400
CTGCTGCTGA AGCCGCATTG GCTTTTGAGG CTGAGAAGAG TGATGCGACG	450
ATGCAAAATG GCCTGAACAT GGAGGAGACG ACGGCGGCGG CTTCTCAGAC	500
TGAAGTGAGT GACACGACGA CAGATCATGG CATGAACATG GAGGAGACAA	550
CGGCGGTGGC TTCTCAGGCT GAGGTGAATG ACACGACGAC AGATCATGGC	600
GTAGACATGG AGGAGACGAT GGTGGAGGCT GTTTTTACTG AGGAACAAAG	650
TGAAGGGTTC AACATGGCGA AGGAGTCGAC GGCGGAGGCT GCTGTTGTTA	700
CGGAGGAACT GAGCAAAGGA GTTTACATGG ACGAGGAGTG GACGTACGAG	750
ATGCCGACCT TGTTGGCTGA TATGGCGGCA GGGATGCTTT TGCCGCCACC	800
ATCTGTACAA TGGGGACATA ATGATGACTT GGAAGGAGAT GCGGACATGA	850
ACCTACTGGA GTTATTAAGG ATCCGCG	877

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 287
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica oleracea
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A

(G) CELL TYPE: N/A  
 (H) CELL LINE: N/A  
 (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
 (A) NAME/KEY: BOCBF5-PEP  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Met	Asn	Thr	Phe	Pro	Ala	Ser	Thr	Glu	Met	Val	Ser	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Tyr	Pro	Met	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Leu	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Met	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Tyr	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	His	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	Ala	Thr	Met	Gln	Asn	Gly	Leu	140	145	150
Asn	Met	Glu	Glu	Thr	Thr	Ala	Ala	Ala	Ser	Gln	Thr	Glu	Val	Ser	155	160	165
Asp	Thr	Thr	Thr	Asp	His	Gly	Met	Asn	Met	Glu	Glu	Thr	Thr	Ala	170	175	180
Val	Ala	Ser	Gln	Ala	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	185	190	195
Val	Asp	Met	Glu	Glu	Thr	Met	Val	Glu	Ala	Val	Phe	Thr	Glu	Glu	200	205	210
Gln	Ser	Glu	Gly	Phe	Asn	Met	Ala	Lys	Glu	Ser	Thr	Ala	Glu	Ala	215	220	225
Ala	Val	Val	Thr	Glu	Glu	Leu	Ser	Lys	Gly	Val	Tyr	Met	Asp	Glu	230	235	240
Glu	Trp	Thr	Tyr	Glu	Met	Pro	Thr	Leu	Leu	Ala	Asp	Met	Ala	Ala	245	250	255
Gly	Met	Leu	Leu	Pro	Pro	Pro	Ser	Val	Gln	Trp	Gly	His	Asn	Asp	260	265	270
Asp	Leu	Glu	Gly	Asp	Ala	Asp	Met	Asn	Leu	Leu	Glu	Leu	Leu	Arg	275	280	285
Ile	Arg																

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 374
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: brCBF1 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CATCCCATTT ACAGGGGGGT TCGTTTAAGA AAGTCAGGTA AGTGGGTGTG	50
TGAAGTGAGG GAACCAAACA AGAAATCTAG GATTGCGTC GGAACTTTCA	100
AAACCGCTGA GATCGCTGCT CGTGCTCAG ACGTTGCTGC CTTAGCCCTC	150
CGCGGGAGAG GCGCCTGCCT CAACTTCGCC GACTCGGCTT GGCGGCTCCG	200
TATCCCGGAG ACAACCTGCG CCAAGGACAT CCAGAAGGCG GCTGCTGAAG	250
CTGCATTGGC TTTTGAGGCC GAGAAGAGTG ATCATGGCAT GAACATCAAG	300
AATACTACGG CGGTGGTTTC TCAGGTTGAG GTGAATGACA CGACGACGGA	350
CCACGGCTTG GACATGGAGG AGAC	374

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Brassica rapa
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY: BRCBF1-PEP
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	Trp	5	10	15
Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	Leu	20	25	30
Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	Val	35	40	45
Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	Leu	Asn	Phe	Ala	50	55	60
Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	Ala	Lys	65	70	75
Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Glu	Ala	80	85	90
Glu	Lys	Ser	Asp	His	Gly	Met	Asn	Ile	Lys	Asn	Thr	Thr	Ala	Val	95	100	105
Val	Ser	Gln	Val	Glu	Val	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Leu	110	115	120
Asp	Met	Glu	Glu												125		

## (2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 884
  - (B) TYPE: Nucleic Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Brassica rapa*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: brCBF2 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

TACACTCAGC CTTATCCAGT TTTTTC AAA AGACTTTTCA ACAATGAACA	50
CATTCCCTGC GTCCACTGAA ATGGTTGGCT CCGAGAACGA GTCTCCGGTT	100
ACTACGGTAG CAGGAGGTGA TTATTATCCC ATGTTGGCGG CAAGCTGTCC	150
GAAGAAGCCA GCGGGTAGGA AGAAGTTTCA GGAGACACGT CACCCCATTT	200
ACCGAGGAGT TCGTCTGAGA AAGTCAGGTA AGTGGGTGTG TGAAGTGAGG	250
GAACCAAACA AGAAATCTAG AATTTGGCTT GGAAC TTTCA AAACAGCTGA	300
GATGGCAGCT CGTGCTCACG ACGTCGCTGC CCTAGCCCTC CGTGGAAGAG	350
GCGCCTGCCT CAATTATGCG GACTCGGCTT GGCGGCTCCG CATCCCGGAG	400
ACAACCTGCC ACAAGGATAT CCAGAAGGCT GCTGCTGAAG CCGCATTGGC	450
TTTTGAGGCT GAGAAAAGTG ATGTGACGAT GCAAAATGGC CTGAACATGG	500
AGGAGATGAC GGCGGTGGCT TCTCAGGCTG AAGTGAATGA CACGACGACA	550
GAACATGGCA TGAACATGGA GGAGGCAACG GCAGTGGCTT CTCAGGCTGA	600
GGTGAATGAC ACGACGACGG ATCATGGCGT AGACATGGAG GAGACAATGG	650
TGGAGGCTGT TTTTACTGAG GAACAAAGTG AAGGGTTTAA CATGGCGAAG	700
GAGTCGACGG TGGAGGCTGC TGTTGTTACG GAGGAACCGA GCAAAGGATC	750
TTACATGGAC GAGGAGTGGA TGCTCGAGAT GCCGACCTTG TTGGCTGATA	800
TGGCGGAAGG GATGCTTTTG CCGCCGCCGT CCGTACAATG GGGACAGAAT	850

GATGACTTCG AAGGAGATGC TGACATGAAC CTCT

884

## (2) INFORMATION FOR SEQ ID NO:77:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 280
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: BRCBF2-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

```

Met Asn Thr Phe Pro Ala Ser Thr Glu Met Val Gly Ser Glu Asn
      5              10              15
Glu Ser Pro Val Thr Thr Val Ala Gly Gly Asp Tyr Tyr Pro Met
      20              25              30
Leu Ala Ala Ser Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe
      35              40              45
Gln Glu Thr Arg His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys
      50              55              60
Ser Gly Lys Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser
      65              70              75
Arg Ile Trp Leu Gly Thr Phe Lys Thr Ala Glu Met Ala Ala Arg
      80              85              90
Ala His Asp Val Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Cys
      95              100             105
Leu Asn Tyr Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr
      110             115             120
Thr Cys His Lys Asp Ile Gln Lys Ala Ala Glu Ala Ala Leu
      125             130             135
Ala Phe Glu Ala Glu Lys Ser Asp Val Thr Met Gln Asn Gly Leu

```

76/101

Asn Met Glu Glu	140	Met Thr Ala Val Ala	145	Ser Gln Ala Glu Val	150
Asp Thr Thr Thr	155	Glu His Gly Met Asn	160	Met Glu Glu Ala Thr	165
Val Ala Ser Gln	170	Ala Glu Val Asn Asp	175	Thr Thr Thr Asp His	180
Val Asp Met Glu	185	Glu Thr Met Val Glu	190	Ala Val Phe Thr Glu	195
Gln Ser Glu Gly	200	Phe Asn Met Ala Lys	205	Glu Ser Thr Val Glu	210
Ala Val Val Thr	215	Glu Glu Pro Ser Lys	220	Gly Ser Tyr Met Asp	225
Glu Trp Met Leu	230	Glu Met Pro Thr Leu	235	Leu Ala Asp Met Ala	240
Gly Met Leu Leu	245	Pro Pro Pro Ser Val	250	Gln Trp Gly Gln Asn	255
Asp Phe Glu Gly	260	Asp Ala Asp Met Asn	265	Leu	270
	275		280		

## (2) INFORMATION FOR SEQ ID NO:78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 806
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: brCBF3 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

ACACTCAGCC TTATCCAGTT TTCAAAAAA GTATTCAACG ATGAACTCAG

50

TCTCTACTTT TTCTGAACTG CTCTGCTCCG AGAACGAGTC TCCGGTTAAT	100
ACGGAAGGTG GTGATTACAT TTTGGCGGCG AGCTGTCCCA AGAAACCTGC	150
TGGTAGGAAG AAGTTTCAGG AGACACGCCA CCCCATTTAC AGAGGAGTTC	200
GTCTGAGGAA GTCAGGTAAG TGGGTGTGTG AAGTGAGGGA ACCAAACAAG	250
AAATCTAGAA TTTGGCTCGG AACTTTCAAA ACAGCTGAGA TCGCAGCTCG	300
TGCTCACGAC GTTGCCGCCT TAGCTCTCCG TGGAAGAGGC GCCTGCCTCA	350
ACTTCGCCGA CTCGGCTTGG CGGCTCCGTA TCCCGGAGAC GACCTGCGCC	400
AAGGATATCC AGAAGGCTGC TGCTGAAGCC GCATTGGCTT TTGAGGCCGA	450
GAAGAGTGAT ACCACGACGA ATGATCGTGG CATGAACATG GAGGAGACGT	500
CGGCGGTGGC TTCTCCGGCT GAGTTGAATG ATACGACGGC GGATCATGGC	550
CTGGACATGG AGGAGACGAT GGTGGAGGCT GTTTTTAGGG AGGAACAGAG	600
AGAAGGGTTT TACATGGCGG AGGAGACGAC GGTGGAGGGT GTTGTTCGG	650
AGTAACAGAT GAGCAAAGGG TTTTACATGG ACGAGGAGTG GACGTTCGAG	700
ATGCCGAGGT TGTGGCTGA TATGGCGGAA GGGATGCTTT TGCCGCCCCC	750
GTCCGTACAA TGGGGACATA ACGATGACTT CGAAGGAGAT GCTGACATGA	800
ACCTCT	806

## (2) INFORMATION FOR SEQ ID NO:79:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 204
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A



- (ix) FEATURE:  
 (A) NAME/KEY: BRCBF3-PEP  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

- (x) PUBLICATION INFORMATION:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

```

Met Asn Ser Val Ser Thr Phe Ser Glu Leu Leu Cys Ser Glu Asn
      5      10      15
Glu Ser Pro Val Asn Thr Glu Gly Gly Asp Tyr Ile Leu Ala Ala
      20      25      30
Ser Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Gln Glu Thr
      35      40      45
Arg His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys
      50      55      60
Trp Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp
      65      70      75
Leu Gly Thr Phe Lys Thr Ala Glu Ile Ala Ala Arg Ala His Asp
      80      85      90
Val Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Cys Leu Asn Phe
      95     100     105
Ala Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Ala
     110     115     120
Lys Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Glu
     125     130     135
Ala Glu Lys Ser Asp Thr Thr Thr Asn Asp Arg Gly Met Asn Met
     140     145     150
Glu Glu Thr Ser Ala Val Ala Ser Pro Ala Glu Leu Asn Asp Thr
     155     160     165
Thr Ala Asp His Gly Leu Asp Met Glu Glu Thr Met Val Glu Ala
     170     175     180
Val Phe Arg Glu Glu Gln Arg Glu Gly Phe Tyr Met Ala Glu Glu
     185     190     195
Thr Thr Val Glu Gly Val Val Pro Glu
      200      205

```

- (2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 755  
 (B) TYPE: Nucleic Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Brassica rapa  
 (B) STRAIN:  
 (C) INDIVIDUAL ISOLATE: N/A  
 (D) DEVELOPMENTAL STAGE: N/A  
 (E) HAPLOTYPE: N/A  
 (F) TISSUE TYPE: N/A

(G) CELL TYPE: N/A  
 (H) CELL LINE: N/A  
 (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
 (A) NAME/KEY: brCBF4 gene  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

ACCGCTCGAG TACTTACTAT ACTACACTCA GCCTTATCCA GTTTTCTTC	50
CAACGATGGA CTCAATCTCT ACTTTTCCTG AACTGCTTGG CTCAGAGAAC	100
GAGTCTCCGG TTACTACGGT AGTAGGAGGT GATTATTGTC CCAGGTTGGC	150
GGCAAGCTGT CCGAAGAAGC CAGCGGGTAG GAAGAAGTTT CAGGAGACAC	200
GTCACCCCAT TTACCGTGGA GTTCGTTTAA GAAAGTCCGG TAAGTGGGTG	250
TGTGAAGTGA GGGAAACAAA CAAGAAATCT AGGATTGTC TCGGAACTTT	300
CAAAACCGCT GAGATCGCTG CTCGTGCTCA CGACGTTGCT GCCTTAGCCC	350
TCCGCGGAAG AGGCGCCTGC CTCAACTTCG CCGACTCGGC TTGACGGCTC	400
CGTATCCCGG AGACAACCTG CGCCAAGGAT ATCCAGAAGG CTGCTGCTGA	450
AGCTGCATTG GCTTTTGAGG CCGAGAAGAG TGATCATGGC ATGAACATGA	500
AGAATACTAC GGCGGTGGCT TCTCAGGTTG AGGTGAATGA TACGACGACG	550
GACCATGGCG TGGACATGGA GGAGACGAGG GTGGAGGGTG TTGTTACGGA	600
GGAACAGAAC AATTGGTTTT ACATGGACGA GGAGTGGATG TTTGGGATGC	650
CGACGTTGTT GGTTGATATG GCGGAAGGGA TGCTTATACC GCGGCAGTCC	700
GTACAATCGG GACACTACGA TGACTTCGAA GGAGATGCTG ACATGAACCT	750
CTGGA	755

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 112  
 (B) TYPE: Amino Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: BRCBF4-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Met	Asp	Ser	Ile	Ser	Thr	Phe	Pro	Glu	Leu	Leu	Gly	Ser	Glu	Asn	
				5					10					15	
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Cys	Pro	Arg	
				20					25					30	
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	
				35					40					45	
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	
				50					55					60	
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	
				65					70					75	
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	
				80					85					90	
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	
				95					100					105	
Leu	Asn	Phe	Ala	Asp	Ser	Ala									
															110

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 832
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa

(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY: brCBF5 gene  
(B) LOCATION:  
(C) IDENTIFICATION METHOD: sequencing  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:  
ACCGCTCGAG TACTTACTAT ACTACACTCA GCCTTATCCA GTTTTTCTTC 50  
CAACGATGGA CTCAATCTCT ACTTTTCCTG AACTGCTTGG CTCAGAGAAC 100  
GAGTCTCCGG TTACTACGGT AGTAGGAGGT GATTATTGTC CCAGGTGGC 150  
GGCAAGCTGT CCGAAGAAGC CAGCGGGTAG GAAGAAGTTT CAGGAGACAC 200  
GTCACCCCAT TTACCGTGGA GTTCGTTTAA GAAAGTCCGG TAAGTGGGTG 250  
TGTGAAGTGA GGGAACCAAA CAAGAAATCT AGGATTTGGC TCGGAACTTT 300  
CAAAACCGCT GAGATCGCTG CTCGTGCTCA CGACGTTGCT GCCTTAGCCC 350  
TCCGCGGAAG AGGCGCCTGC CTCAACTTCG CCGACTCGGC TTGGCGGCTC 400  
CGTATCCCGG AGACAACCTG CGCCAAGGAT ATCCAGAAGG CTGCTGCTGA 450  
AGCTGCTTTG GCTTTTGAGG CCGAGAAGAG TGATCATGGC ATGAACATGA 500  
AGAATACTAC GGCGGTGGCT TCTCAGGTTG AGGTGAATGA TACGACGACG 550  
GACCATGGCG TGGACATGGA GGAGACGTTG GTGGAGGCTG TTTTACGGA 600  
GGAACAGAGA GAAGGGTTTT ACATGACGGA GGAGACGAGG GTGGAGGGTG 650  
TTGTTACGGA GGAACAGAAC AATTGGTTTT ACATGGACGA GGAGTGGATG 700  
TTTGGGATGC CGACGTTGTT GGTGATATG GCGGAAGGGA TGCTTATACC 750  
GCGGCAGTCC GTACAATCGG GACACTACGA TGACTTCGAA GGAGATGCTG 800  
ACATGAACCT CTGGAATTAT TAGGGATCCG CG 832

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 255  
(B) TYPE: Amino Acid  
(C) STRANDEDNESS: Single

- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Protein
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Brassica rapa
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY: BRCBF5-PEP
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Met	Asp	Ser	Ile	Ser	Thr	Phe	Pro	Glu	Leu	Leu	Gly	Ser	Glu	Asn	5	10	15
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Cys	Pro	Arg	20	25	30
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	35	40	45
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	50	55	60
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	65	70	75
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	80	85	90
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	95	100	105
Leu	Asn	Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	110	115	120
Thr	Cys	Ala	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	125	130	135
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	His	Gly	Met	Asn	Met	Lys	Asn	140	145	150
Thr	Thr	Ala	Val	Ala	Ser	Gln	Val	Glu	Val	Asn	Asp	Thr	Thr	Thr	155	160	165
Asp	His	Gly	Val	Asp	Met	Glu	Glu	Thr	Leu	Val	Glu	Ala	Val	Phe	170	175	180
Thr	Glu	Glu	Gln	Arg	Glu	Gly	Phe	Tyr	Met	Thr	Glu	Glu	Thr	Arg	185	190	195
Val	Glu	Gly	Val	Val	Thr	Glu	Glu	Gln	Asn	Asn	Trp	Phe	Tyr	Met			

	200		205		210
Asp Glu Glu Trp	Met Phe Gly Met Pro	Thr Leu Leu Val Asp	Met		
	215		220		225
Ala Glu Gly Met	Leu Ile Pro Arg Gln	Ser Val Gln Ser Gly	His		
	230		235		240
Tyr Asp Asp Phe	Glu Gly Asp Ala Asp	Met Asn Leu Trp Asn	Tyr		
	245		250		255

## (2) INFORMATION FOR SEQ ID NO:84:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 830
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: brCBF6 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TACTACACTC AGCCTTATCC AGTTTTCAAA AAAAGTATTC AACTATGAAC	50
TCAGTCTCTA CTTTTTCTGA ACTGCTCTGC TCCGAGAACA AGTCTCCGGT	100
TAATACGGAA GGTGGTGATT ACATTTTGGC GGCGAGCTGT CCCAAGAAAC	150
CTGCTGGTAG GAAGAAGTTT CAGGAGACAC GCCACCCCAT TTACAGAGGA	200
GTTCGCCTAA GAAAGTCAGG TAAGTGGGTG TGTGAAGTGA GGGAAACCAA	250
CAAGAAATCT AGAATTTGGC TCGGAACTTT CAAAACAGCT GAGATAGCAG	300
CTCGTGCTCA CGACGTCGCC GCCTTAGCTC TCCGTGGAAG AGGCGCCTGC	350

CTCAACTTCG CCGACTCGGC TTGGCGGCTC CGTATCCCAG AGACAACCTG	400
CGCCAAGGAT ATCCAGAAGG CTGCTGCTGA AGCCGCATTG GCTTTTGAGG	450
CCGAGAAGAG TGATACCACG ACGAATGATC GTGGCATGAA CATGGAGGAG	500
ACGTCCGCGG TGGCTTCTCC GGCTGAGTTG AATGATACGA CGGCGGATCA	550
TGGCCTGGAC ATGGAGGAGA CGATGGTGGA GGCTGTTTTT AGGGACGAAC	600
AGAGAGAAGG GTTTTACATG GCGGAGGAGA CGACGGTGGA GGGTGTGTGTT	650
CCGGAGGAAC AGATGAGCAA AGGGTTTTAC ATGGACGAGG AGTGGACGTT	700
CGAGATGCCG AGGTTGTTGG CTGATATGGC GGAAGGGATG CTTCTGCCTC	750
CCCCGTCCGT ACAATGGGGA CATAACGATG ACTTCGAAGG AGATGCTGAC	800
ATGAACCTCT GGAATTATTA GGGATCCGCG	830

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY: BRCBF6-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

Met	Asn	Ser	Val	Ser	Thr	Phe	Ser	Glu	Leu	Leu	Cys	Ser	Glu	Asn	
				5					10					15	
Lys	Ser	Pro	Val	Asn	Thr	Glu	Gly	Gly	Asp	Tyr	Ile	Leu	Ala	Ala	
				20					25					30	
Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Gln	Glu	Thr	
				35					40					45	
Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	Ser	Gly	Lys	
				50					55					60	
Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	Trp	
				65					70					75	
Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	Asp	
				80					85					90	
Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	Leu	Asn	Phe	
				95					100					105	
Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	Ala	
				110					115					120	
Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	Ala	Phe	Glu	
				125					130					135	
Ala	Glu	Lys	Ser	Asp	Thr	Thr	Thr	Asn	Asp	Arg	Gly	Met	Asn	Met	
				140					145					150	
Glu	Glu	Thr	Ser	Ala	Val	Ala	Ser	Pro	Ala	Glu	Leu	Asn	Asp	Thr	
				155					160					165	
Thr	Ala	Asp	His	Gly	Leu	Asp	Met	Glu	Glu	Thr	Met	Val	Glu	Ala	
				170					175					180	
Val	Phe	Arg	Asp	Glu	Gln	Arg	Glu	Gly	Phe	Tyr	Met	Ala	Glu	Glu	
				185					190					195	
Thr	Thr	Val	Glu	Gly	Val	Val	Pro	Glu	Glu	Gln	Met	Ser	Lys	Gly	
				200					205					210	
Phe	Tyr	Met	Asp	Glu	Glu	Trp	Thr	Phe	Glu	Met	Pro	Arg	Leu	Leu	
				215					220					225	
Ala	Asp	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Pro	Pro	Ser	Val	Gln	
				230					235					240	
Trp	Gly	His	Asn	Asp	Asp	Phe	Glu	Gly	Asp	Ala	Asp	Met	Asn	Leu	
				245					250					255	
Trp	Asn	Tyr													

## (2) INFORMATION FOR SEQ ID NO:86:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 854
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Brassica rapa
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A



(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: brCBF7 gene

(B) LOCATION:

(C) IDENTIFICATION METHOD: sequencing

(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CTATACTACA CACAGCCTTA TCCAGCCGCT CGAGTACTTA CTATACTACA	50
CTCAGCCTTT TCCAGTTTTT CAAAAGAAGT TTTCAACGAT GAACTCAGTC	100
TCTACTCTTT CTGAAGTTCT TGGCTCCCAG AACGAGTCTC CCGTAGGTGG	150
TGATTACTGT CCCATGTTGG CGGCGAGCTG TCCGAAGAAG CCGGCGGGTA	200
GGAAGAAGTT TCGGGAGACA CGTCACCCCA TTTACAGAGG AGTTCGTCTT	250
AGAAAGTCAG GTAAGTGGGT GTGTGAAGTG AGGGAACCAA ACAAGAAATC	300
TAGGATTTGG CTCGGAACCT TCAAAACAGC TGAGATCGCA GCTCGTGCTC	350
ACGACGTTGC CGCCTTAGCT CTCCGTGGAA GAGGCGCCTG CCTCAACTTC	400
GGCGACTCGG CTTGGCGGCT CCGTATCCCG GAGACAACCT GCGCCAAGGA	450
TATCCAGAAG GCTGCTGCTG AAGCCGCATT GGCTTTTGAG GCGGAGAAGA	500
GTGATACCAC GACGACGAAT GATCATGGCA TGAACATGGC TTCTCAGGTT	550
GAGGTTAATG ACACGACGGA TCATGACCTG GACATGGAGG AGACGATGGT	600
GGAGGCTGTT TTTAGGGAGG AACAGAGAGA AGGGTTTTAC ATGGCGGAGG	650
AGACGACGGT GGAGGGTATT GTTCCGGAGG AACAGATGAG CAAAGGGTTT	700
TACATGGACG AGGAGTGGAT GTTCGGGATG CCGACCTTGT TGGCTGATAT	750
GGCGGCAGGG ATGCTCTTAC CGCCGCCGTC CGTACAATGG GGACATAATG	800
ATGACTTCGA AGGAGATGCT GACATGAACC TCTGGAATTA TTAAGGGATC	850
CGCG	854

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 251

(B) TYPE: Amino Acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: Brassica rapa
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
- (A) NAME/KEY: BRCBF7-PEP
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

```

Met Asn Ser Val Ser Thr Leu Ser Glu Val Leu Gly Ser Gln Asn
      5              10              15
Glu Ser Pro Val Gly Gly Asp Tyr Cys Pro Met Leu Ala Ala Ser
      20              25              30
Cys Pro Lys Lys Pro Ala Gly Arg Lys Lys Phe Arg Glu Thr Arg
      35              40              45
His Pro Ile Tyr Arg Gly Val Arg Leu Arg Lys Ser Gly Lys Trp
      50              55              60
Val Cys Glu Val Arg Glu Pro Asn Lys Lys Ser Arg Ile Trp Leu
      65              70              75
Gly Thr Phe Lys Thr Ala Glu Ile Ala Ala Arg Ala His Asp Val
      80              85              90
Ala Ala Leu Ala Leu Arg Gly Arg Gly Ala Cys Leu Asn Phe Ala
      95              100             105
Asp Ser Ala Trp Arg Leu Arg Ile Pro Glu Thr Thr Cys Ala Lys
      110             115             120
Asp Ile Gln Lys Ala Ala Ala Glu Ala Ala Leu Ala Phe Glu Ala
      125             130             135
Glu Lys Ser Asp Thr Thr Thr Thr Asn Asp His Gly Met Asn Met
      140             145             150
Ala Ser Gln Val Glu Val Asn Asp Thr Thr Asp His Asp Leu Asp
      155             160             165
Met Glu Glu Thr Met Val Glu Ala Val Phe Arg Glu Glu Gln Arg
      170             175             180
Glu Gly Phe Tyr Met Ala Glu Glu Thr Thr Val Glu Gly Ile Val
      185             190             195
Pro Glu Glu Gln Met Ser Lys Gly Phe Tyr Met Asp Glu Glu Trp
      200             205             210
Met Phe Gly Met Pro Thr Leu Leu Ala Asp Met Ala Ala Gly Met
      215             220             225
Leu Leu Pro Pro Pro Ser Val Gln Trp Gly His Asn Asp Asp Phe

```

	230		235	240
Glu Gly Asp Ala	Asp Met Asn Leu Trp	Asn Tyr		
	245	250		

## (2) INFORMATION FOR SEQ ID NO:88:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 738
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Glycine max
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: gmCBF1 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

CATCCGATTT ATAGTGGCGT GAGGAGGAGG AACACGGATA AGTGGGTAAG	50
TGAGGTGAGG GAGCCCAACA AAAAGACCAG GATTTGGCTG GGGACTTTTC	100
CCACGCCGGA GATGGCGGCA CGGGCCACG ACGTGGCGGC AATGGCCCTG	150
AGGGGCCGGT ATGCCTGTCT CAACTTCGCT GACTCGACGT GGC GGTTACC	200
AATCCCGCC ACTGCTAACG CAAAGGATAT ACAGAAAGCA GCAGCAGAGG	250
CTGCCGAGGC TTTCAGACCA AGTCAGACCT TAGAAAATAC GAATACAAAG	300
CAAGAGTGTG TAAAAGTGGT GACGACAACA ACGATCACAG AACAAAAACG	350
AGGAATGTTT TATACGGAGG AAGAAGAGCA AGTGTTAGAT ATGCCTGAGT	400
TGCTTAGGAA TATGGTGCTT ATGTCCCAA CACATTGCAT AGGGTATGAG	450

```

TATGAAGATG CTGACTTGGG TGCTCAAGAT GCTGAGGTGT CCCTATGGAG      500
TTTCTCAATT TAATAACGTG CTTTGGTTT GGTTTTTTAT GTTAGTTTTG      550
GAGTGTGACT GTCTGTACTG GTTTTTTATT AGTAGTACGG ATACTAGCTA      600
TAGGTGGCAG ATTGAAAGGG ACCAAAAGGA ATTTTCTTTT GAAACCCTTT      650
TTGTCAAAGT AATCAATCGC GTATCATCAA GTGAATCCCT TGATCAAGTT      700
TATGTATGAA TTAAATAAAA GAAGAATCTA GTTTTGGT                    738

```

## (2) INFORMATION FOR SEQ ID NO:89:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Protein

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Glycine max
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: GMCBF1-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

```

His Pro Ile Tyr Ser Gly Val Arg Arg Arg Asn Thr Asp Lys Trp
              5              10              15
Val Ser Glu Val Arg Glu Pro Asn Lys Lys Thr Arg Ile Trp Leu
              20              25              30
Gly Thr Phe Pro Thr Pro Glu Met Ala Ala Arg Ala His Asp Val
              35              40              45
Ala Ala Met Ala Leu Arg Gly Arg Tyr Ala Cys Leu Asn Phe Ala
              50              55              60

```

Asp	Ser	Thr	Trp	Arg	Leu	Pro	Ile	Pro	Ala	Thr	Ala	Asn	Ala	Lys	
				65					70					75	
Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Glu	Ala	Phe	Arg	Pro	
				80					85					90	
Ser	Gln	Thr	Leu	Glu	Asn	Thr	Asn	Thr	Lys	Gln	Glu	Cys	Val	Lys	
				95					100					105	
Val	Val	Thr	Thr	Thr	Thr	Ile	Thr	Glu	Gln	Lys	Arg	Gly	Met	Phe	
				110					115					120	
Tyr	Thr	Glu	Glu	Glu	Gln	Val	Leu	Asp	Met	Pro	Glu	Leu	Leu		
				125					130					135	
Arg	Asn	Met	Val	Leu	Met	Ser	Pro	Thr	His	Cys	Ile	Gly	Tyr	Glu	
				140					145					150	
Tyr	Glu	Asp	Ala	Asp	Leu	Asp	Ala	Gln	Asp	Ala	Glu	Val	Ser	Leu	
				155					160					175	
Trp	Ser	Phe	Ser	Ile											
				170											

## (2) INFORMATION FOR SEQ ID NO:90:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 793
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Raphanus sativus*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

## (vii) IMMEDIATE SOURCE: N/A

## (viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: rsCBF1 gene
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

ACTACACTCA GCCTTATCCA GTTTTCTTC CAACGATGGA CTCAATCTCT	50
ACTTTTCTG AACTGCTTGG CTCCGAGAAC GAGTCTCCGG TTACTACGGT	100
AGTAGGAGGT GATTATTTTC CCAGGTTGGC GGCAAGCTGT CCGAAGAAGC	150

CAGCGGGTAG GAAGAAGTTT CAGGAGACAC GTCACCCCAT TTACCGCGGA	200
GTTCGTTTAA GAAAGTCAGG TAAGTGGGTG TGTGAAGTGA GGAACCAAAA	250
CAAGAAATCT AGGATTTGGC TCGGAAC TTT CAAAACCGCT GAGATCGCTG	300
CTCGTGCTCA CGACGTTGCT GCCTTAGCCC TCCGCGGAAG AGGCGCCTGC	350
CTCAACTTCG CCGACTCGGC TTGGCGGCTC CGTATCCCGG AGACAACCTG	400
CGCCAAGGAT ATCCAGAAGG CTGCTGCTGA AGCTGCATTG GCTTTTGAGG	450
CCGAGAAGAG TGATCATGGC ATGAACATGA AGAATACTAC GGCGGTGGCT	500
TCTCAGGTTG AGGTGAATGA CACGACGACG GACCATGGCG TGGACATGGA	550
GGAGACGTTG GTGGAGGCTG TTTTACGGA GGAACAGAGA GAAGGGTTTT	600
ACATGACGGA GGAGACGAGG GTGGAGGGTG TTGTTACGGA GGAACAGAAC	650
AATTGGTTTT ACATGGACGA GGAGTGGATG TTTGGGATGC CGACGTTGTT	700
GGTTGATATG GCGGAAGGGA TGCTTTTACC GCGGCCGTCC GTACAATCGG	750
GACACTACGA TGA CTTCGAA GGAGATGCTG ACATGAACCT CTG	793

## (2) INFORMATION FOR SEQ ID NO:91:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 252
- (B) TYPE: Amino Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Raphanus sativus*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: RSCBF1-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing

## (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Met	Asp	Ser	Ile	Ser	Thr	Phe	Ser	Glu	Leu	Leu	Gly	Ser	Glu	Asn	
				5					10					15	
Glu	Ser	Pro	Val	Thr	Thr	Val	Val	Gly	Gly	Asp	Tyr	Phe	Pro	Arg	
				20					25					30	
Leu	Ala	Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	
				35					40					45	
Gln	Glu	Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Lys	
				50					55					60	
Ser	Gly	Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	
				65					70					75	
Arg	Ile	Trp	Leu	Gly	Thr	Phe	Lys	Thr	Ala	Glu	Ile	Ala	Ala	Arg	
				80					85					90	
Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Gly	Ala	Cys	
				95					100					105	
Leu	Asn	Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	
				110					115					120	
Thr	Cys	Ala	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Leu	
				125					130					135	
Ala	Phe	Glu	Ala	Glu	Lys	Ser	Asp	His	Gly	Met	Asn	Met	Lys	Asn	
				140					145					150	
Thr	Thr	Ala	Val	Ala	Ser	Gln	Val	Glu	Val	Asn	Asp	Thr	Thr	Thr	
				155					160					165	
Asp	His	Gly	Val	Asp	Met	Glu	Glu	Thr	Leu	Val	Glu	Ala	Val	Phe	
				170					175					180	
Thr	Glu	Glu	Gln	Arg	Glu	Gly	Phe	Tyr	Met	Thr	Glu	Glu	Thr	Arg	
				185					190					195	
Val	Glu	Gly	Val	Val	Thr	Glu	Glu	Gln	Asn	Asn	Trp	Phe	Tyr	Met	
				200					205					210	
Asp	Glu	Glu	Trp	Met	Phe	Gly	Met	Pro	Thr	Leu	Leu	Val	Asp	Met	
				215					220					225	
Ala	Glu	Gly	Met	Leu	Leu	Pro	Arg	Pro	Ser	Val	Gln	Ser	Gly	His	
				230					235					240	
Tyr	Asp	Asp	Phe	Glu	Gly	Asp	Ala	Asp	Met	Asn	Leu				
				245					250						

## (2) INFORMATION FOR SEQ ID NO:92:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 682
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (iii) HYPOTHETICAL: No

## (iv) ANTI-SENSE: No

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Raphanus sativus
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A

(E) HAPLOTYPE: N/A  
 (F) TISSUE TYPE: N/A  
 (G) CELL TYPE: N/A  
 (H) CELL LINE: N/A  
 (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

(A) NAME/KEY: rsCBF2 gene  
 (B) LOCATION:  
 (C) IDENTIFICATION METHOD: sequencing  
 (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

ACACCTAAAC CTTATCCAGG TTAACTTTT TTTTCATAA AGAGTTTTCA	50
ACAATGACCA CATTCTCTAC CTTTCCGAA ATGTTGGGCT CCGAGTACGA	100
GTCTCCGGTT ACATTAGGCG GAGAGTATTG TCCGAAGCTG GCCGCGAGCT	150
GTCCGAAGAA ACCAGCTGGT CGTAAGAAGT TTCGAGAGAC GCGCCACCCA	200
ATATACAGAG GAGTTCGTCT GAGAACTCA GGTAAGTGGG TGTGTGAAGT	250
GAGGGAGCCA AACAAGAAAT CTAGGATTTG GCTCGGTACT TTCCTAACCG	300
CCGAGATCGC AGCGCGTGCC CACGACGTCG CCGCCATAGC CCTCCGCGGC	350
AAATCCGCAT GTCTCAATTT CGCTGACTCG GCTTGGCGGC TCCGTATCCC	400
GGAGACAACA TGCCCCAAGG ATATACAGAA GGCGGCTGCT GAAGCCGCGG	450
TGGCTTTTCA GGCTGAGATA AATGATACGA CGACGGATCA TGGCCTGGAC	500
TTGGAGGAGA CGATCGTGGA GGCTATTTTT ACGGAGGTAA ACAACGATGA	550
GTTTTATATG GACGAGGAGT CCATGTTTCG GATGCCGTCT TTGTTGGCTA	600
GTATGGCGGA AGGGATGCTT TTGCCGCTGC CGTCCGTACA ATCTGAACAT	650
AACTGTGACT TCGACGGAGA TGCTGACATG AA	682

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 209  
 (B) TYPE: Amino Acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:



## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Raphanus sativus*
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: RSCBF2-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Met	Thr	Thr	Phe	Ser	Thr	Phe	Ser	Glu	Met	Leu	Gly	Ser	Glu	Tyr	5	10	15
Glu	Ser	Pro	Val	Thr	Leu	Gly	Gly	Glu	Tyr	Cys	Pro	Lys	Leu	Ala	20	25	30
Ala	Ser	Cys	Pro	Lys	Lys	Pro	Ala	Gly	Arg	Lys	Lys	Phe	Arg	Glu	35	40	45
Thr	Arg	His	Pro	Ile	Tyr	Arg	Gly	Val	Arg	Leu	Arg	Asn	Ser	Gly	50	55	60
Lys	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys	Ser	Arg	Ile	65	70	75
Trp	Leu	Gly	Thr	Phe	Leu	Thr	Ala	Glu	Ile	Ala	Ala	Arg	Ala	His	80	85	90
Asp	Val	Ala	Ala	Ile	Ala	Leu	Arg	Gly	Lys	Ser	Ala	Cys	Leu	Asn	95	100	105
Phe	Ala	Asp	Ser	Ala	Trp	Arg	Leu	Arg	Ile	Pro	Glu	Thr	Thr	Cys	110	115	120
Pro	Lys	Asp	Ile	Gln	Lys	Ala	Ala	Ala	Glu	Ala	Ala	Val	Ala	Phe	125	130	135
Gln	Ala	Glu	Ile	Asn	Asp	Thr	Thr	Thr	Asp	His	Gly	Leu	Asp	Leu	140	145	150
Glu	Glu	Thr	Ile	Val	Glu	Ala	Ile	Phe	Thr	Glu	Val	Asn	Asn	Asp	155	160	165
Glu	Phe	Tyr	Met	Asp	Glu	Glu	Ser	Met	Phe	Gly	Met	Pro	Ser	Leu	170	175	180
Leu	Ala	Ser	Met	Ala	Glu	Gly	Met	Leu	Leu	Pro	Leu	Pro	Ser	Val	185	190	195
Gln	Ser	Glu	His	Asn	Cys	Asp	Phe	Asp	Gly	Asp	Ala	Asp	Met		200	205	

## (2) INFORMATION FOR SEQ ID NO:94:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 349
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Zea maize
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY: zmCBF1 gene
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD: sequencing
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

GGAGTCCGC GGACGGCGGC GGCGGCGGCG ACGACGAGTA CGCGACGGTG	50
CTGTCGGCGC CACCCAAGCG GCCGGCGGGG CGGACCAAGT TCCGGGAGAC	100
GCGGCACCCC GTGTACCGCG GCGTGCGGCG GCGCGGGCCC GCGGGGCGCT	150
GGGTGTGCGA GGTCCGCGAG CCAACAAGA AGTCGCGCAT CTGGCTCGGC	200
ACCTTCGCCA CCCCCGAGGC CGCCGCGCGC GCGCACGACG TGGCCGCGCT	250
GGCCCTGCGG GGCCGCGCCG CGTGCCTCAA CTTCGCCGAC TCGGCGCGCC	300
TGCTCCAAGT CGACCCCGCC ACGCTCGCCA CCCCCGACGA CATCCGCCG	349

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 115
  - (B) TYPE: Amino Acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Protein
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Zea maize
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

## (ix) FEATURE:

- (A) NAME/KEY: ZMCBF1-PEP
- (B) LOCATION:
- (C) IDENTIFICATION METHOD: sequencing
- (D) OTHER INFORMATION:

## (x) PUBLICATION INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Glu	Ser	Ala	Asp	Gly	Gly	Gly	Gly	Gly	Asp	Asp	Glu	Tyr	Ala	Thr		
				5					10					15		
Val	Leu	Ser	Ala	Pro	Pro	Lys	Arg	Pro	Ala	Gly	Arg	Thr	Lys	Phe		
				20					25					30		
Arg	Glu	Thr	Arg	His	Pro	Val	Tyr	Arg	Gly	Val	Arg	Arg	Arg	Gly		
				35					40					45		
Pro	Ala	Gly	Arg	Trp	Val	Cys	Glu	Val	Arg	Glu	Pro	Asn	Lys	Lys		
				50					55					60		
Ser	Arg	Ile	Trp	Leu	Gly	Thr	Phe	Ala	Thr	Pro	Glu	Ala	Ala	Ala		
				65					70					75		
Arg	Ala	His	Asp	Val	Ala	Ala	Leu	Ala	Leu	Arg	Gly	Arg	Ala	Ala		
				80					85					90		
Cys	Leu	Asn	Phe	Ala	Asp	Ser	Ala	Arg	Leu	Leu	Gln	Val	Asp	Pro		
				95					100					105		
Ala	Thr	Leu	Ala	Thr	Pro	Asp	Asp	Ile	Arg							
				110					115							

## (2) INFORMATION FOR SEQ ID NO:96:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:

(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GGAAGATCTA TGAAACAGAG TACTCTGATC AATGAACTC

39

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 37  
(B) TYPE: Nucleic Acid Sequence  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:  
(A) ORGANISM:  
(B) STRAIN:  
(C) INDIVIDUAL ISOLATE: N/A  
(D) DEVELOPMENTAL STAGE: N/A  
(E) HAPLOTYPE: N/A  
(F) TISSUE TYPE: N/A  
(G) CELL TYPE: N/A  
(H) CELL LINE: N/A  
(I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A  
(viii) POSITION IN GENOME: N/A

(ix) FEATURE:  
(A) NAME/KEY:  
(B) LOCATION:  
(C) IDENTIFICATION METHOD:  
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

98/101

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

GGAAGATCTG AACAGAGTA CTCTGATCAA TGAACTC

37

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- (B) STRAIN:
- (C) INDIVIDUAL ISOLATE: N/A
- (D) DEVELOPMENTAL STAGE: N/A
- (E) HAPLOTYPE: N/A
- (F) TISSUE TYPE: N/A
- (G) CELL TYPE: N/A
- (H) CELL LINE: N/A
- (I) ORGANELLE: N/A

(vii) IMMEDIATE SOURCE: N/A

(viii) POSITION IN GENOME: N/A

(ix) FEATURE:

- (A) NAME/KEY:
- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GGAAGATCTA TGAACAGAGT ACTCTGATCA ATGAACTC

38

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39
- (B) TYPE: Nucleic Acid Sequence
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:
  - (B) LOCATION:
  - (C) IDENTIFICATION METHOD:
  - (D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGAAGATCTA TGAACAGAGT ACTCTGATGC AATGAACTC

39

(2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 49
  - (B) TYPE: Nucleic Acid Sequence
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (iii) HYPOTHETICAL: No
- (iv) ANTI-SENSE: No
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM:
  - (B) STRAIN:
  - (C) INDIVIDUAL ISOLATE: N/A
  - (D) DEVELOPMENTAL STAGE: N/A
  - (E) HAPLOTYPE: N/A
  - (F) TISSUE TYPE: N/A
  - (G) CELL TYPE: N/A
  - (H) CELL LINE: N/A
  - (I) ORGANELLE: N/A
- (vii) IMMEDIATE SOURCE: N/A
- (viii) POSITION IN GENOME: N/A
- (ix) FEATURE:
  - (A) NAME/KEY:

100/101

- (B) LOCATION:
- (C) IDENTIFICATION METHOD:
- (D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GGAGGATCCT CGTTTCTACA ACAATAAAAT AAAATAAAAT GAAGGAACC

49



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup>:</b>  <b>C12N 15/29, C07K 14/415, C12N 5/10, A01H 5/00, A01N 65/00</b>	<b>A3</b>	<b>(11) International Publication Number:</b> <b>WO 99/38977</b>  <b>(43) International Publication Date:</b> 5 August 1999 (05.08.99)																																	
<b>(21) International Application Number:</b> PCT/US99/01895  <b>(22) International Filing Date:</b> 28 January 1999 (28.01.99)  <b>(30) Priority Data:</b> <table style="width: 100%; border: none;"> <tr> <td style="width: 30%;">09/018,233</td> <td style="width: 40%;">3 February 1998 (03.02.98)</td> <td style="width: 30%;">US</td> </tr> <tr> <td>09/017,816</td> <td>3 February 1998 (03.02.98)</td> <td>US</td> </tr> <tr> <td>09/018,235</td> <td>3 February 1998 (03.02.98)</td> <td>US</td> </tr> <tr> <td>09/017,575</td> <td>3 February 1998 (03.02.98)</td> <td>US</td> </tr> <tr> <td>09/018,227</td> <td>3 February 1998 (03.02.98)</td> <td>US</td> </tr> <tr> <td>09/018,234</td> <td>3 February 1998 (03.02.98)</td> <td>US</td> </tr> <tr> <td>09/198,119</td> <td>23 November 1998 (23.11.98)</td> <td>US</td> </tr> </table>		09/018,233	3 February 1998 (03.02.98)	US	09/017,816	3 February 1998 (03.02.98)	US	09/018,235	3 February 1998 (03.02.98)	US	09/017,575	3 February 1998 (03.02.98)	US	09/018,227	3 February 1998 (03.02.98)	US	09/018,234	3 February 1998 (03.02.98)	US	09/198,119	23 November 1998 (23.11.98)	US	Filed on 3 February 1998 (03.02.98) US 09/198,119 (CIP) Filed on 23 November 1998 (23.11.98)												
09/018,233	3 February 1998 (03.02.98)	US																																	
09/017,816	3 February 1998 (03.02.98)	US																																	
09/018,235	3 February 1998 (03.02.98)	US																																	
09/017,575	3 February 1998 (03.02.98)	US																																	
09/018,227	3 February 1998 (03.02.98)	US																																	
09/018,234	3 February 1998 (03.02.98)	US																																	
09/198,119	23 November 1998 (23.11.98)	US																																	
<b>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications</b> <table style="width: 100%; border: none;"> <tr> <td style="width: 30%;">US</td> <td style="width: 40%;">09/018,233 (CIP)</td> <td style="width: 30%;"></td> </tr> <tr> <td>Filed on</td> <td>3 February 1998 (03.02.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/017,816 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>3 February 1998 (03.02.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/018,235 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>3 February 1998 (03.02.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/017,575 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>3 February 1998 (03.02.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/018,227 (CIP)</td> <td></td> </tr> <tr> <td>Filed on</td> <td>3 February 1998 (03.02.98)</td> <td></td> </tr> <tr> <td>US</td> <td>09/018,234 (CIP)</td> <td></td> </tr> </table>		US	09/018,233 (CIP)		Filed on	3 February 1998 (03.02.98)		US	09/017,816 (CIP)		Filed on	3 February 1998 (03.02.98)		US	09/018,235 (CIP)		Filed on	3 February 1998 (03.02.98)		US	09/017,575 (CIP)		Filed on	3 February 1998 (03.02.98)		US	09/018,227 (CIP)		Filed on	3 February 1998 (03.02.98)		US	09/018,234 (CIP)		<b>(71) Applicants (for all designated States except US):</b> MENDEL BIOTECHNOLOGY, INC. [US/US]; 21375 Cabot Boulevard, Hayward, CA 94545 (US). MICHIGAN STATE UNIVERSITY [US/US]; 238 Administration Building, East Lansing, MI 48824-1046 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> STOCKINGER, Eric, J. [US/US]; 1360 Burcham Drive, East Lansing, MI 48823 (US). JAGLO-OTTOSEN, Kirsten [US/US]; 307 South Clemens Avenue, Lansing, MI 48912 (US). ZARKA, Daniel [US/US]; 2101 Barritt Street, Lansing, MI 48912 (US). GILMOUR, Sarah, Jane [GB/US]; 1830 Barnes Road, Leslie, MI 49251 (US). JIANG, Cai-Zhong [CN/US]; 34495 Heathrow Terrace, Fremont, CA 94555 (US). FROMM, Michael [US/US]; 968 Keeler Avenue, Berkeley, CA 94708 (US). THOMASHOW, Michael, F. [US/US]; 805 Southlawn Avenue, East Lansing, MI 48823 (US).  <b>(74) Agent:</b> GUERRERO, Karen., J.; MENDEL BIOTECHNOLOGY, INC., 21375 Cabot Boulevard, Hayward, CA 94545 (US).  <b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>  <b>(88) Date of publication of the international search report:</b> 28 October 1999 (28.10.99)
US	09/018,233 (CIP)																																		
Filed on	3 February 1998 (03.02.98)																																		
US	09/017,816 (CIP)																																		
Filed on	3 February 1998 (03.02.98)																																		
US	09/018,235 (CIP)																																		
Filed on	3 February 1998 (03.02.98)																																		
US	09/017,575 (CIP)																																		
Filed on	3 February 1998 (03.02.98)																																		
US	09/018,227 (CIP)																																		
Filed on	3 February 1998 (03.02.98)																																		
US	09/018,234 (CIP)																																		
<b>(54) Title:</b> PLANT HAVING ALTERED ENVIRONMENTAL STRESS TOLERANCE																																			
<b>(57) Abstract</b>  A transformed plant is provided which comprises one or more environmental stress tolerance genes; a DNA regulatory sequence which regulates expression of the one or more environmental stress tolerance genes; a sequence encoding a binding protein capable of binding to the DNA regulatory sequence and inducing expression of the one or more environmental stress tolerance genes; and a recombinant promoter which regulates expression of the gene encoding the binding protein. A method for altering an environmental stress tolerance of a plant is also provided which comprises the steps of transforming a plant with a promoter which regulates expression of at least one copy of a gene encoding a binding protein capable of binding to a DNA regulatory sequence which regulates one or more environmental stress tolerance genes in the plant; expressing the binding protein encoded by the gene; and stimulating expression of at least one environmental stress tolerance gene through binding of the binding protein to the DNA regulatory sequence.																																			



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## 1. Claims: 1-30 all partially

A binding protein other than CBF-1 in isolated form comprising a consensus sequence sufficiently homologous to or being the consensus sequence shown in Figure 19A, 19B, or 19C, that the protein is capable of binding to a CCG regulatory sequence. Said protein comprising a sequence selected from the group consisting of an AP2 domain of SEQ ID Nos. 13 and 15.

A non-naturally occurring binding protein comprising an amino acid sequence capable of binding to a CCG regulatory sequence and an amino acid sequence which forms a transcription activation region.

DNA encoding said proteins, said proteins recombinantly expressed within a plant, nucleic acid constructs and expression constructs comprising said DNA and a promoter being tissue specific, e.g. flower, fruit, or seed specific in plants, as well as plant material comprising said subject matter.

A method for altering an environmental stress tolerance of a plant by expressing said proteins in a plant and thereby stimulating the expression of at least one environmental stress tolerance gene through binding of said binding protein to a DNA regulatory sequence comprising CCG which regulates the expression of one or more environmental stress tolerance genes in the plant, optionally expressing the protein at levels and under conditions at which the plant does not express the binding protein in the plant's native state.

## 2. Claims: 1-30 all partially

idem for SEQ ID NOs: 39,41,43,45

## 3. Claims: 1-30 all partially

idem for SEQ ID NOs:47,49,51,53,55,57,59,61,63

## 4. Claims: 1-30 all partially

idem for SEQ ID NOs:65,67,69,71,73

## 5. Claims: 1-30 all partially

idem for SEQ ID NOs:75,77,79,81,83,85,87

## 6. Claims: 1-30 all partially

idem for SEQ ID NO:89

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

7. Claims: 1-30 all partially

idem for SEQ ID NOs:91,93

8. Claims: 1-30 all partially

idem for SEQ ID NO:95

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/01895

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/29 C12N5/10 C07K14/415 A01H5/00 A01N65/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A01H A01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>BÜTTNER M. AND SINGH K.: "Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein"</p> <p>PNAS, U.S.A.,</p> <p>vol. 94, no. 11, May 1997 (1997-05), pages 5961-5966, XP002108953</p> <p>the whole document</p> <p style="text-align: center;">---</p> <p style="text-align: center;">-/--</p>	1-3,8,9

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents:

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*Z\* document member of the same patent family

Date of the actual completion of the international search

13 July 1999

Date of mailing of the international search report

03.09.99

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Kania, T

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/JS 99/01895

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OHME-TAKAGI M. AND SHINSHI H.: "Ethylene-inucible DNA binding proteins that interact with an ethylene-responsive element" THE PLANT CELL, vol. 7, no. 2, 1995, pages 173-182, XP002108954 cited in the application the whole document	1-3,8,9
X	--- WILSON K. ET AL.: "A dissociation insertion causes a semidominant mutation that increases expression of TINY, an Arabidopsis gene related to APETALA2" THE PLANT CELL, vol. 8, no. 4, 1996, pages 659-671, XP002108955 cited in the application the whole document	1-3,6,8, 9
X	--- STOCKINGER E J ET AL: "Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C - repeat / DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA, (1997 FEB 4) 94 (3) 1035-40., XP002108956 cited in the application see esp. p.1038 l.par. - p.1039; discussion	1-9
A	--- CA 2 104 142 A (LALIBERTE JEAN FRANCOIS ;HOUDE MARIO (CA); SARHAN FATHEY (CA)) 17 February 1995 (1995-02-17) see esp. p.23, p.42-44 l.19, p.50	1-30
A	--- OKAMURO J. ET AL.: "The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis" PNAS,U.S.A., vol. 94, no. 13, June 1997 (1997-06), pages 7076-7081, XP002108957 cited in the application the whole document	1-30
	--- -/--	

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/01895

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	SHINWARI Z K ET AL: "An Arabidopsis gene family encoding DRE /CRT binding proteins involved in low-temperature-responsive gene expression." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (1998 SEP 8) 250 (1) 161-70., XP002108958 see esp. figure 5; discussion ---	1-30
P,X	GILMOUR S J ET AL: "Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold -induced COR gene expression." PLANT JOURNAL, (1998 NOV) 16 (4) 433-42., XP002108959 the whole document ---	1-30
P,X	MEDINA J. ET AL.: "The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration" PLANT PHYSIOLOGY, vol. 119, no. 2, February 1999 (1999-02), pages 463-469, XP002108960 the whole document ---	1-30
T	WO 98 09521 A (UNIV MICHIGAN) 12 March 1998 (1998-03-12) the whole document -----	1-30

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 99/01895

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
  
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-30, all partially.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JS 99/01895

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CA 2104142 A	17-02-1995	US 5731419 A	24-03-1998
WO 9809521 A	12-03-1998	US 5892009 A	06-04-1999
		AU 4157797 A	26-03-1998
		US 5891859 A	06-04-1999